LUX 2.7 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.NullAccessControlProcessor`, `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	positive integer, default 50

	Correlated Rule will hold in memory per bucket, for example per unique attribute	
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

matching.engine.streamanalyzer.sampling.enabled	Whether the stream analyzer should analyze every event, or sample. Sampling is first 1000, then every 5th for 15 minutes, then one every 15 minutes. Default true.	boolean
matching.engine.streamanalyzer.ignore.props	Comma separated list of event property prefixed to ignore. E.g. “track_point_”	Comma separated list
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
operator.buffer.size	Sets the size of event buffers between operators	default 100
event.ingest.fail.fast	If true, engine will start even if some Event Ingesters fail to initialize	Default false

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
```

1.1.5 Track Manager Configuration

Configure the track manager per stream to enable NAI Enters, Exits, and Crosses operators.

```
track.manager.max.track.length=4
track.manager.id.query.stream.1=wmata_stream
track.manager.id.query.path.1=attributes/vehicleId
track.manager.geo.query.stream.1=wmata_stream
track.manager.geo.query.path.1=/geometries
```

1.1.6 Admin Console Configuration

Configure the Admin Console to enable the engine to send status updates and receive commands from the AdminConsole webapp. Defaults:

```
#turn the admin console webpage updates on or off
admin.console.web.enabled=true
#Update frequency for REST calls, updates
admin.console.update.frequency.ms=10000
#Base URL for admin console REST services
admin.console.rest.url=https://localhost/AdminConsole/rest/
admin.console.rest.session.id=PleaseChangeMe

#When enabling the file service, additional config changes must be made in the
AdminConsole .war's config or it will throw exceptions
#admin.console.file.service.enabled=true
#admin.console.file.service.exclusions=.standalone_lock
#admin.console.file.service.inclusions=

#Be especially careful about enabling writes in the file service. It shouldn't
be used unless additional security
#measures are in place (e.g. the webapp is not publicly accessible and/or
certs are required at the container level)
#admin.console.file.service.write.enabled=false
#admin.console.file.service.write.exclusions=logs,bin,lib,ext/lib,plugins,gc.l
og,.standalone_lock,*.jar,*.sh
#admin.console.file.service.write.inclusions=
```

```
#admin.console.file.service.file.limit.directory=1000
#admin.console.file.service.file.limit.total=10000
#admin.console.file.service.file.poll.interval=20000
#admin.console.file.service.request.poll.interval=5000
```

1.2 analytics.xml

`analytics.xml` controls which Analytic plugins are loaded by the engine. The following table lists options that can be configured in `analytics.xml`.

Property	Description	Data Type or Valid Values
analytics_path	Disk location of Analytic plugins	Valid folder path
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
analytic_file (multiple)	Path to file containing analytic definition to load.	File path

analytics.xml options

1.2.1 Analytic plugin file

Each `analytic_file` entry in `analytics.xml` should point to a file containing an analytic plugin definition. The following fields are required in addition to properties specific to the analytic, listed in Section 2.1.

stream_in	Specifies a data stream to send to the analytic	A stream name specified in an Ingest plugin's configuration
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 this plugin's name and configuration	A valid classification string
description	Description sent to User Interface	string
display_name	Analytic name for User Interface and Admin Console	string
overlay	Whether an Analytic plugin outputs a map layer for the User Interface	boolean

Sample analytic xml file:

```
<analytic
name="icg.engine.analytic.normalcy.entitygeo.EntityGeospatialNormalcyAnalytic"
```

```
classification="UNCLASSIFIED" overlay="false"
stream_out="entity_geospatial_normalcy_analytic_java" display_name="Entity
Geospatial Normalcy">
    <stream_in>myStream</stream_in>
    <description>Determines the geospatial normalcy for an event</description>
    <deadlock_check_time_millis>10000</deadlock_check_time_millis>
    <property name="id_attribute_xpath" value="/attributes/mmsi"/>
    …
   <property name="training_period_s" value="0"/>
</analytic>
```

1.3 enrichments.xml

`enrichments.xml` controls which Enrichment plugins are loaded by the engine. The following table lists options that can be configured in `enrichments.xml`.

Property	Description	Data Type or Valid Values
enrichments_path	Disk location of Enrichment plugins	Valid folder path
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
enrichment_file (multiple)	Path to file containing enrichment definition to load.	File path

enrichments.xml options

1.3.1 Enrichment plugin file

Each `enrichment_file` entry in `enrichments.xml` should point to a file containing an enrichment plugin definition. The following fields are required in addition to properties specific to the enrichment, listed in Section 2.2.

stream_in	Specifies a data stream to send to the enrichment	A stream name specified in an Ingest plugin's configuration
name	The Java class name containing the plugin code	A class name on the runtime classpath
classification	Classification of this plugin's name and configuration	A valid classification string
description	Description sent to User Interface	string
display_name	Enrichment name for User Interface and Admin Console	string

1.4 event_outputs.xml

`event_outputs.xml` controls which Event Output plugins are loaded by the engine. The following table lists options that can be configured in `event_outputs.xml`.

Property	Description	Data Type or Valid Values
plugin_path	Disk location of Event Outpout plugins	Valid folder path
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
event_output_file (multiple)	Path to file containing event output definition to load.	File path

event_outputs.xml options

1.4.1 Event Output plugin file

Each `event_output_file` entry in `event_outputs.xml` should point to a file containing an event output plugin definition. The following fields are required in addition to properties specific to the event output, listed in Section 2.4.

stream_in	Specifies a data stream to send to the event output	A stream name specified in an Ingest plugin's configuration
name	The Java class name containing the plugin code	A class name on the runtime classpath
classification	Classification of this plugin's name and configuration	A valid classification string
description	Description sent to User Interface	string
display_name	Event Output name for User Interface and Admin Console	string

1.4.2 Enabling Event Outputs

Within `/bin/setenv.sh` alter `ENGINE_OPTS` to include event output by modifying the line to equal `ENGINE_OPTS="--include-event-output"`.

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	Postive long or hh:mm:ss string
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 a region, in meters. 0 (default) means the event geo overlaps or is contained by the region geo.	long
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	boolean

	max_distance_to_region_m	

Table 4 - GeoAnalyticFilter options

Some analytics define `grid_size_m` for geohash grid sizes. Though they take arbitrary value, geohashes can only be in the following sizes (which are themselves approximations, as actual grid size varies with latitude). Values entered will be rounded up to the closest valid size.
`5003500, 625400, 123260, 19540, 3800, 605, 116, 18, 3, 0.5`

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 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 or hh:mm:ss string
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 or hh:mm:ss string
track.classifier.x	Optional, up to 3 names of	"shape", "icg", or "kmeans"

	TrackClassifiers to load.	
<filter>	GeoAnalyticFilter	
track_style_name	Style name for track, default **`abnormal_track`**	String
point_style_name	Style name for track points, default **`abnormal_track_point`**	String
shape.x.name	For “shape” track classifier	String
shape.x.file	For “shape” track classifier	String
min_match_threshold	For “shape” track classifier	double
add_rotations	For “shape” track classifier	boolean
kmeans_pmml_model_file	For “kmeans” track classifier	File path
write_to_graph_store	Whether or not to write entities and relationships to the Entity Manager Graph Store	boolean
<EntityConfig properties>	If write_to_graph_store is true	

Table 5 - Abnormal Track Analytic options

Output Property	Description	Data Type
numPosits	The number of events in the abnormal track	int
travelDistance	The distance travelled in the abnormal track	Double
distanceDelta	The distance between the first and last points in the track	Double
percentage	distanceDelta / travelDistance	Double
description	Description of result, including id of entity making abnormal track	String
id	Unique id of entity	String
track_point_x	Geo for event x in track	GeoJSON point
event_time_x	Time for event x in track	long
entity_id	If entity manager is enabled, ID of entity from entity database	

Abnormal Track Analytic output properties

Output Property	Description	Data Type
analytic_abnormal_track_location	Location of abnormal track first position	GEO
analytic_abnormal_track_time	Time of abnormal track detection	DATESTRING
analytic_abormal_track_descriptio n	Display name of analytic	String

<Entity>	If it doesnt already exist, entity will be created	

Abnormal Track Analytic EM Entity properties

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 long or hh:mm:ss string
window_size_seconds	Amount of data to hold in memory, measured as seconds since current time	positive long or hh:mm:ss string
write_to_graph_store	Whether or not to write entities and relationships to the Entity Manager	boolean

	Graph Store	
<filter>	EventFilter	
aggregate_classifications	Whether to accumulate and combine classifications and data groups	boolean
<RelationshipConfig properties>	If write_to_graph_store is true	

Table 6 - Association Analytic options

Output Property	Description	Data Type
description	Description of result, including id of source and dest entities	String
source	ID of source entity	String
destination	ID of destination entity	String
relationship_id	If entity manager is enabled, ID of relationship from entity database	String

Association Analytic output properties

Output Property	Description	Data Type
<Relationship>		

Association Analytic EM properties

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 (proximity).	positive long
max_association_time_s	Max temporal delta between events for IDs to be associated, in seconds	positive long or hh:mm:ss string

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 an attribute required to be present in the event	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 time	positive long or hh:mm:ss string
<filter>	GeoFilter	
track_1_style_name	Default `track`, style name for track 1 polylines	String
track_2_style_name	Default `track`, style name for track 2 polylines	String
write_to_graph_store	Whether or not to write entities and relationships to the Entity Manager Graph Store	boolean
<RelationshipConfig properties>	If write_to_graph_store is true	

Table 7 - CoTravel Analytic options

Output Property	Description	Data Type
description	Description of result, including id of source and dest entities	String
total_events	Number of events that generated this analytic event	Int
travel_distance1	Distance travelled by entity 1 in the co-travel track	Double
travel_distance2	Distance travelled by entity 2 in the co-travel track	Double
uniqueAttribute1	ID of entity 1	
uniqueAttribute2	ID of entity 2	
requiredAttribute1	Required attribute of entity 1	
requiredAttribute2	Required attribute of entity 2	
track_1	Track for entity 1	GeoJSON polyline
track_2	Track for entity 2	GeoJSON polyline
relationship_id	If entity manager is enabled, ID of relationship from entity database	String

CoTravel output properties

Output Property	Description	Data Type
<Relationship>		

CoTravel Analytic EM properties

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 or hh:mm:ss string
<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.

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	
prediction_track_style_name	Geo style name for predicted path, default **`predicted_path`**	String
prediction_point_style_name	Geo style name for predicted point, default **`predicted_point`**	String
prediction_size_s	How many seconds into the future to predict	positive long or hh:mm:ss string

Table 9 - Dead Reckoning Analytic options

Output Property	Description	Data Type
nextLat	Predicted latitude	double
nextLon	Predicted longitude	double
nextTime	ETA at predicted lat/lon	String
lastKnownLat	Last known latitude	Double
lastKnownLon	Last known longitude	double
lastKnownTime	Time at last observation	String
groupingAttribute	ID of entity	String
predictionPath	Line between current and predicted locations	GeoJSON line
predictionPoint	Predicted point	GeoJSON point

Dead Reckoning Analytic output properties

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	Positive long

	can be apart to qualify for an analytic event	
<filter>	GeoAnalyticFilter	
write_to_graph_store	Whether or not to write entities and relationships to the Entity Manager Graph Store	boolean
<EntityConfig properties>	If write_to_graph_store is true	

Table 10 - Dupe ID Analytic options

Output Property	Description	Data Type
description	Description of result, including dupe ID	String
id	ID of the dupe entity	String
distance	Distance in meters between the two observations	double
time	Time delta in seconds between the two observations	long
travelSpeed	Required travel speed for a single entity to be responsible for both observations, in meters per second	double
entity_id	If entity manager is enabled, ID of entity from entity database	String

Dupe ID output properties

Output Property	Description	Data Type
analytic_dupe_id_time	Time of dupe detection	DATESTRING
analytic_dupe_id_location	Location of dupe detection	GEO
analytic_dupe_id_description	Display name of analytic	String

Dupe ID Analytic EM Entity properties

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	5003500, 625400, 123260, 19540, 3800, 605, 116, 18, 3, 0.5
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	Postive long or hh:mm:ss string
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
area_style_name	Style name for map in UI, default `normalcy_area`	String

Table 11 - Entity Geospatial Normalcy Analytic options

Output Property	Description	Data Type
value	The number of observations	int
mean	The mean number of observations for that area	double
stdDev	The standard deviation for observations in that area	double
geoHash	The geoHash area	String
geoBounds	Rectangular geometry around the area	GeoJSON polygon
normalcyScore	0 (abnormal) to 1 (normal) score for this analytic event	double

Entity Geospatial Normalcy Analytic output properties

EntityGeospatialNormalcyAnalytic can be trained via the standalone class `icg.engine.analytic.normalcy.entitygeo.EGNATrainer.`. It takes 1 command line

argument which is the path to a config file. The config file should have all parameters for the analytic specified, and a `data_folder` parameter, which is the location of csv input files for the analytic to train on. You can also, optionally, specify testing points to see if the track will produce output based on the trained file. The format of the csv file should be: `eventId,timestamp,entity id,lat,lon.` `eventId` is not currently used, so the first column can be anything. Sample config file:

```
geom_query=/event/georss:where
grid_size_m=2000
id_attribute_xpath=/event/id
id_attribute_regex=.*
history_file_name=egna.dat
data_folder=/mnt/hgfs/Shared/nm4/csv
use_month=false
use_day=false
use_hour=true
use_minute=false
use_day_of_week=false
use_second=false
normalcy_threshold=.5

test_point.1=67.214,-20.808
test_id.1=710005960
```

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 time	Positive long or hh:mm:ss string
<filter>	GeoAnalyticFilter	
cluster_style_name	Style name for cluster bounding circle, default `cluster`	String
write_to_graph_store	Whether or not to write entities and relationships to the Entity Manager Graph Store	boolean
<RelationshipConfig properties>	If write_to_graph_store is true	

Table 13 - Geohash Clustering Analytic options

Output Property	Description	Data Type
cluster_area	Geo bounds of cluster	GeoJSON circle
clusterCenter	Center geo for cluster	GeoJSON point
clusterCenterLatitude	Latitude of center of cluster	double
clusterCenterLongitude	Longitude of center of cluster	double
diameter	Diameter of cluster in meters	double
totalDistanceMeters	Total distance between all events in the cluster, in meters	double
clusterStartTime	Time of earliest event in the cluster	long
clusterDurationMs	Time delta between earliest and latests cluster events, in milliseconds	long
common_attribute	The common attribute value for events in the cluster, if set	String
description	Description string including common attribute	String
number_unique_ids	Number of unique ids in the cluster, if configured	int
unique_ids	Each unique ID in the cluster	String
relationship_id	If entity manager is enabled, ID of relationship from entity database	String

Geohash Clustering Analytic output properties

Output Property	Description	Data Type
<Relationship>	For each pair in the cluster	

Geohash Clustering Analytic EM Entity properties

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	Postive long or hh:mm:ss string
<filter>	GeoAnalyticFilter	
association_line_style_name	Style name for line between associated events, default `association_line`	String
write_to_graph_store	Whether or not to write entities and relationships to the Entity Manager Graph Store	boolean
<RelationshipConfig properties>	If write_to_graph_store is true	

Table 14 - Geohash Proximity Analytic options

Output Property	Description	Data Type
distance	Distance between the entities in meters	double
description	Description string including both IDs	String
id_1	The ID of entity 1	String
id_2	The ID of entity 2	String
proximity	Geometry connecting two positions	GeoJSON line
relationship_id	If entity manager is enabled, ID of relationship from entity database	String

Geohash Proximity Analytic output properties

Output Property	Description	Data Type
<Relationship>		

Geohash Proximity Analytic EM properties

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	positive integer

	traverse to meet min_nodes_for_cluster nodes	
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	Postive long or hh:mm:ss string
window_size_seconds	Amount of data to hold in memory, measured as seconds since current time	Postive long or hh:mm:ss string
<filter>	EventFilter	
aggregate_classifications	Whether to accumulate and combine classifications and data groups	boolean
write_to_graph_store	Whether or not to write entities and relationships to the Entity Manager Graph Store	boolean
<RelationshipConfig properties>	If write_to_graph_store is true	

Table 15 - Graph Clusters Analytic options

Output Property	Description	Data Type
description	Description of the cluster including node IDs	String
window	Window_size in hh:mm:ss	string
clusterNode	One property for each node in the cluster, value will be the ID	String
edgeProperty	If any edge attributes are configured, hey will be output here	String
connections_<node id>	Comma separated list of connections for the node	String
clusterImage	JPG data for rendering cluster image	Binary image
relationship_id	If entity manager is enabled, ID of relationship from entity database	String

Graph Clusters output properties

Output Property	Description	Data Type
<Relationship>	For each pair in the cluster	

Graph Clusters Analytic EM properties

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 or hh:mm:ss string
<filter>	GeoAnalyticFilter postfix “A” GeoAnalyticFilter postfix “B”	One for each stream
association_line_style_name	Style name for line between associated events, default `association_line`	String
write_to_graph_store	Whether or not to write entities and relationships to the Entity Manager Graph Store	boolean
<RelationshipConfig properties>	If write_to_graph_store is true	

Table 16 - Group DAOI Analytic options

Output Property	Description	Data Type
description	Description as configured	String
distance	Distance between stream A and B events	double

id	The ID of the entity in the stream A event	String
closest_event	Line between event closest stream A event to the stream B event	GeoJSON line

Group DAOI output properties

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	5003500, 625400, 123260, 19540, 3800, 605, 116, 18, 3, 0.5
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_include_itemlist	Whether to include a list of items in the kml descriptions	boolean
kml_output_file	Optional file to output KML to	File path string
kml_output_interval_s	How often to output KML, in seconds	Postive long or hh:mm:ss string
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	Postive long or hh:mm:ss string

<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 or hh:mm:ss string
<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 or hh:mm:ss string
check_interval_s	How often to check timeout window, in seconds	Positive long or hh:mm:ss string
missing_value_display_name	Display name for missing IDs	String

Table 18 - ID Inactivity Analytic options

Output Property	Description	Data Type
geo	Geo of last event	GeoRSS XML
age_s	Age of last event in seconds	long
missing_value	The inactive ID	String
last_seen	Date of last observation	Date string

ID Inactivity Analytic output properties

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	Postive long or hh:mm:ss string
<filter>	GeoAnalyticFilter	

Table 19 - Moving Analytic options

Output Property	Description	Data Type
travelDistance	Distance travelled in meters	double
description	Description including entity ID	String
id	ID of the entity	String

Moving Analytic output properties

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 list will be ignored.	File path

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	5003500, 625400, 123260, 19540, 3800, 605, 116, 18, 3, 0.5
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 in a 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, ROLLING_NORMAL_STAT_WITH_HI	Integer >= 0

	STORY. 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
include_incomplete_buckets	If set to false, when the state is loaded, counts for the time bucket in which the state was saved as well as the time bucket in which it was loaded will be ignored. (default true)	boolean
realtime_lag_limit_millis	If set to a positive integer, the analytic will wait this long after the end of a time bucket for more events to come in	Integer >= 0

	with eventTime values in that bucket before closing it and outputting. Note that the bucket will still be closed immediately if an event comes in with an eventTime after its end.	
time_lag_limit_millis	Similar to realtime_lag_limit_millis, but the bucket will not be closed by out-of-order events as long as they are within the specified time limit. Temporary buckets are created for events after the current bucket end but before the lag limit and copied into the correct bucket for that eventTime once that bucket becomes the current one.	Integer >= 0
drop_late_events	If set to true, events coming in after their time window has ended will be dropped instead of adding to the current window. (default true)	boolean
image_encoder_threads	Number of additional threads to use for creating and encoding graph displays	Integer >= 0
pruner.nonzero_samples_min	Only used with SLIDING_WINDOW_NORMAL_STAT. If specified, this is the number of nonzero entries a fully filled-out window would need to have for a given time bucket for it to 'count' as non-empty for the purposes of pruner.nonempty_timebuckets_min	Integer >= 0
pruner.nonempty_timebuckets_min	Only used with SLIDING_WINDOW_NORMAL_STAT. If specified, this is the number of non-empty time-buckets that must be present for an attribute combination for its data to be retained.	Integer >= 0
<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

Output Property	Description	Data Type
normalcyScore	0 to 1 normalcy score	double
value	Value for current time bucket	double
timeWindow	Current time window	"<date> to <date>" String

geoCenterPoint	Center point of area if confugred	GeoJSON point
normalcyGraph	Graph image	JPG binary
geoBounds	Geohash area bounds if configured	GeoJSON polygon

Multi-Attribute Normalcy output properties

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	Postive long or hh:mm:ss string
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

Output Property	Description	Data Type
description	Description	string

New Value Analytic output properties

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	Postive long or hh:mm:ss string
stop_box_lat_lon	Series of lat/lon coordinate pairs, which represent the corners of stop boxes.	Coordinate string
<filter>	GeoAnalyticFilter	
prediction_track_style_name	Geo style name for predicted path, default `predicted_path`	String
prediction_point_style_name	Geo style name for predicted point, default `predicted_point`	String

Table 23 - Path Projection Analytic options

Output Property	Description	Data Type
lastKnownLat	Latitude of current position	double
lastKnownLon	Longitude of current position	double
lastKnownTime	Time of last observation	Date string
nextLat.x	Latitude of predicted position x	double
nextLon.x	Longitude of predicted position x	double
nextTime.x	Predicted time for position x	Date string
geo.x	Predicted geo for position x	GeoJSON point
predictionPath	Predicted path line through all predicted positions	GeoJSON line

Path Projection Analytic output properties

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. Standard 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 month and this value is 1 or more, you won't get alerts for years.	Integer >= 0
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

Output Property	Description	Data Type
normalcyScore	0 to 1 normalcy score	double
mean	Average value expected	double

stdDev	Standard deviation for values	double
description	Description string for binary mode only	String
value	Observed value	double
id	The entity ID	String
timeWindow	Time window for measurement	“<date> to <date>” string

Pattern of Life Normalcy Analytic output properties

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
check_interval	Interval, in seconds, to check for inactivity.	Positive long or hh:mm:ss string
window_size	Time window size in seconds	Positive long or hh:mm:ss string
<filter>	none	

Table 25 - Stream Inactivity Analytic options

Output Property	Description	Data Type
stream_name	Name of stream	String
last_event	Date of last event	Date string
active	If the stream is active	boolean
went_inactive	Time the stream went inactive	Date string
went_active	Time the stream became active again	Date string
down_time	String down time in minutes	long

Stream Inactivity Analytic output properties

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	Postive long or hh:mm:ss string
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	Comma separated string

	ignore globally	
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 output sub-results	Positive integer
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

Output Property	Description	Data Type
termList	List of top terms	String
dataIntervalDuration	Data interval	Hh:mm String
resultsInterval	Results interval	“<date> to <date>” string
predictionPeriod	Prediction period	“<date> to <date>” string
resultsDuration	Results duration	Hh:mm string
predictedTermList	Predicted next top terms	String
predictionInterval	Interval for prediction	“<date> to <date>” string

Term Frequency Analytic output properties

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	Postive long or hh:mm:ss string
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	
split_spaces_only	If true, only spaces (not punctuation) will be used to split words. Default `false`	Boolean

Table 27 - Term Trend Analytic options

Output Property	Description	Data Type
time_range_start	Start of time range	Date string
time_range_end	End of time range	Date string
previous_time_range_start	Start of previous time interval	Date string

Term Trend Analytic output properties

2.1.24 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	long

	the area geo.	
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.	5003500, 625400, 123260, 19540, 3800, 605, 116, 18, 3, 0.5
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
area_style_name	Style name for map in UI, default `MultiAreaAssociationAnalytic`	String
write_to_graph_store	Whether or not to write entities and relationships to the Entity Manager Graph Store	boolean
<EntityConfig properties>	If write_to_graph_store is true	

Table 28 - Multi Area Association Analytic options

Output Property	Description	Data Type
id	Entity ID	string
description	Description including ID and numAreas	String
areaAssociationName.x	Name of area x	String
areaAssociationObservations.x	Number of observations for area x	int
area_x	Shape for area x	GeoJSON polygon
numAreas	Number of areas entity has been associated with	int
entity_id	If entity manager is enabled, ID of entity from entity database	String

Multi Area Association Analytic output properties

Output Property	Description	Data Type
analytic_multi_area_association_time	Time of analytic event	DATESTRING
analytic_multi_area_association_location	Location of analytic event	GEO

analytic_multi_area_association_description	Display name of analytic	String

Multi Area Association Analytic EM Entity properties

2.1.25 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	Postive long or hh:mm:ss string
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	
track_style_name	Style name for track, default **`track`**	String
point_style_name	Style name for track points, default **`track_point`**	String
add_features	If true, track feature properties will be added to results from	Boolean

	ICGFeatureExtractor. Default **`false`**	
max_events	(Optional) Max events in a track	int
max_travel_distance_meters	(Optional) Max track distance in meters	long

Table 29 - Track Shape Similarity Analytic options

Output Property	Description	Data Type
travelDistance	Track length in meters	double
travelDuration	Time delta between head and tail of track segment, in seconds	double
description	Description including entity ID	String
id	Entity ID	string
track_point_x	Geo for point x in track	GeoJSON point
track	Track line	GeoJSON line

Track Shape Similarity Analytic output properties

2.1.26 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	Postive long or hh:mm:ss string
warning_threshold_s	Time to project tracks into the future for check against warning areas	Postive long or hh:mm:ss string
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	
prediction_track_style_name	Geo style name for predicted path, default **`predicted_path`**	String
write_to_graph_store	Optional, if true will update entities in the EM that are alerted on with a new attribute	boolean
<EntityConfig properties>	If write_to_graph_store is true	

id_xpath	Only if write_to_graph_store is true, xpath to entity ID	XPath string

Table 30 - Area Warning Analytic options

Output Property	Description	Data Type
description	Description including time to area	String
distance_nm	DIstance to area in nautical miles	double
time_s	Time to area in seconds	long
time_hhmmss	Time to area in hh:mm:ss	Hh:mm:ss string
predictionPath	Line for predicted path	GeoJSON line
entity_id	If entity manager is enabled, ID of entity from entity database	String

Area Warning Analytic output properties

Output Property	Description	Data Type
analytic_area_warning_time	Time of analytic event	DATESTRING
analytic_area_warning_location	Location of analytic event	GEO
analytic_area_warning_description	Display name of analytic	String

Multi Area Association Analytic EM Entity properties

2.1.27 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)	Positive long or hh:mm:ss string
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
prediction_track_style_name	Geo style name for predicted path, default **`predicted_path`**	String
prediction_point_style_name	Geo style name for predicted point, default **`predicted_point`**	String

Table 31 - Dead Reckoning Course Speed Analytic options

Output Property	Description	Data Type
nextTime_x	Predicted time at point x	Date string
predictionPoint_x	Geo for predicted point x	GeoJSON point
predictionPath	Line through predicted points	GeoJSON line

Dead Reckoning Course Speed Analytic output properties

2.1.28 Area Pattern of Life Analytic

`icg.engine.analytic.areapol.AreaPOLAnalytic`

The AreaPOL analytic monitors events in a specified AOI, and produces an alert with activity statistics on a configurable time interval defined by `report_interval_s`. The analytic is focused on entity activity within the AOI; entities are defined by `id_attribute_xpath`. The following statistics are included:

Evicted IDs this Period - IDs that timed-out and were removed from statistics this period. The eviction timeout parameter is optional in the analytic configuration.
First Ever Seen IDs - IDs that were seen for the first time ever during this time period.
Departed IDs this Period - IDs that were observed outside of the AOI this period, after previously being seen inside the AOI.
Observations this Period - Total events seen in AOI during this period.
Entities Observed this Period - Unique IDs observed during this period.
Current IDs table - A table containing information for the entities currently within the AOI.
Category counts - Optional. Lists the number of entities belonging to each configured category currently in the AOI.
Top Num Obs table - Lists a configured number of top entities by observation count. Observation count is the number of times the entity has visited the AOI.
Top Total Duration table - Lists a configured number of top entities by total visit duration. Visit duration is only counted if an entity is observed once after it is observed within the AOI and before it is evicted.
Entities in Area by Time graph - Charts the "Entities Observed this Period" statistic
Events in Area by Time graph - Charts the "Observations this Period" statistic

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	Postive long or hh:mm:ss string
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.	Postive long or hh:mm:ss string
<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
area_style_name	Style name for map in UI, default `AreaPOL`	String

Table 32 - Area Pattern of Life Analytic options

Output Property	Description	Data Type
evictedIDs	IDs that were evicted this period	String
firstSeenIDs	First seen IDs this period	String
departingIDs	Departed IDs this period	String
numObs	Number of observations in the area this period	long
numEntities	Number of unique entities observed this period	int
currentId_	ID x currently in area	string
duration_	Duration entity x has been in area	long
firstSeen_	Time entity x first seen in area	Date string
lastSeen_	Time entity x last seen in area	Date string
categoryCount_	Entity count for entity category x	int
area	Region shape	GeoJSON polygon
description	Description including time window	String

Area Pattern of Life Analytic output properties

2.1.29 Geospatial Graph Analytic

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

Constructs a network graph in memory from associations between IDs. IDs are associated if they are in geospatial and temporal proximity as configured. Periodically writes the graph to disk in GraphML and KML formats.

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	Postive long or hh:mm:ss string
history_file_name	File to save/load geospatial map	File path string
graphml_state_file_name	File to save GraphML to	File path string
window_size_s	Time in seconds before an ID is removed from the geospatial map, preventing future associations until it's seen again, default no expire	Postive long or hh:mm:ss string
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
write_to_graph_store	Whether or not to write entities and relationships to the Entity Manager Graph Store	boolean
<RelationshipConfig properties>	If write_to_graph_store is true	

Table 33 - Geospatial Graph Analytic options

GeospatialGraphAnalytic can be trained via the standalone class `icg.engine.analytic.geospatial.graph.GGATrainer`. It takes 1 command line argument which is the path to a config file. The config file should have all parameters for the analytic specified, and a `data_folder` parameter, which is the location of csv input files for the analytic to train on. You can also, optionally, specify testing points to see if the track will produce output based on the trained file. The format of the csv file should be: `eventId,timestamp,entity id,lat,lon.` `eventId` is not currently used, so the first column can be anything. Sample config file:

```
location_xpath=/event/georss:where
id_attribute_xpath=/event/id
graphml_state_file_name=gga.graphml
kml_output_file=gga.kml
history_file_name=gga.dat
distance_threshold_meters=300
```

`data_folder=/mnt/hgfs/Shared/nm4/csv`

Output Property	Description	Data Type
<Relationship>	Relationships for each proximity pair	

Geospatial Graph EM properties

2.1.30 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	5003500, 625400, 123260, 19540, 3800, 605, 116, 18, 3, 0.5
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
area_style_name	Style name for map in UI, default `normalcy_area`	String
point_style_name	Style name for map in UI, default `normalcy_area_point`	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

Table 34 - Geospatial Normalcy Analytic options

Output Property	Description	Data Type
normalcyScore	0 to 1 score of normalcy	double
mean	Average of observed values	double
stdDev	Standard deviation of observed values	double
value	Observed value this period	double
geoHash	Geohash of area	String
timeWindow	Time window of period	“<date> to <date>” string
geoCenterPoint	Center point of area	GeoJSON point
geoBounds	Area shape	GeoJSON polygon

Geospatial Normalcy Analytic output properties

2.1.31 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	Postive long or hh:mm:ss string
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

2.1.32 Geo Grid Track Forecast Analytic

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

Uses historical tracks to create associations between geospatial areas, uses that association graph to predict future points on current tracks. 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	5003500, 625400, 123260, 19540, 3800, 605, 116, 18, 3, 0.5
track_gap_expire_time_s	Max amount of time allowed between track updates before it is considered expired and removed, in seconds	Postive long or hh:mm:ss string
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. Must be greater than or equal to min_geohash_track_size_for_prediction	int

min_geohash_track_size_for_prediction	Minimum number of geohashes present in a track for a prediction to be made, default 2. Must be less than or equal to max_geohash_chain_size.	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	Postive long or hh:mm:ss string
track_style_name	Style name for track line, default **`track`**	String
track_point_style_name	Style name for track point, default **`track_point`**	String
prediction_area_style_name	Style name for prediction square, default **`GeoGridTrackForecast`**	String
prediction_path_style_name	Style name for line between prediction squares, default **`predicted_path`**	String
history_file_name	Optional, file to save learned history to	File path string
report_incorrect_predictions	If true, the analytic will produce events for entities that do not follow their forecasted track. Default false.	Boolean
save_interval_s	Number of seconds between history saves. Optional. Default: **`00:10:00`**	Postive long or hh:mm:ss string
process_out_of_order_updates	If false, old events (out of order) will be ignored. Default **`false`**	Boolean
save_state_only_on_exit	If true, state will not be periodically saved, only on engine shutdown or plugin restart.	Boolean

Table 35 - Geo Grid Track Forecast Analytic options

Output Property	Description	Data Type
description	Description string including ID	String

calc_speed_ms	Calculated average speed in m/s	double
track_length_m	Length of the track in meters	double
track_duration_hhmmss	Duration of the track	Hh:mm:ss string
track_point_x	Geo for point x on track	GeoJSON point
track	Line for track	GeoJSON line
predictionPath	Line for prediction	GeoJSON line
predictionArea	Predicted area (can be more than one)	GeoJSON polygon

Geo Grid Track Forecast Analytic output properties

GeoGridTrackForecastAnalytic can be trained via the standalone class `icg.engine.analytic.track.forecast.GeoGridTrackForecastAnalyticTrainer`. It takes 1 command line argument which is the path to a config file. The config file should have all parameters for the analytic specified, and a `data_folder` parameter, which is the location of csv input files for the analytic to train on. You can also, optionally, specify testing points to see if the track will produce output based on the trained file. The format of the csv file should be: `eventId,timestamp,entity id,lat,lon.` `eventId` is not currently used, so the first column can be anything. Sample config file:

```
location_xpath=/event/georss:where
grid_size_m=2000
id_attribute_xpath=/event/id
max_geohash_chain_size=6
min_geohash_track_size_for_prediction=6
track_gap_expire_time_s=7200
history_file_name=ggtfa.dat
data_folder=/mnt/hgfs/Shared/nm4/csv

test_point.1=67.214,-20.808
test_point.2=67.214,-20.852
test_point.3=67.214,-20.896
test_point.4=67.214,-20.94
test_point.5=67.258,-20.984
test_point.6=67.258,-21.028
```

2.1.33 Simple Normalcy Analytic

`icg.engine.analytic.normalcy.simple.SimpleNormalcyAnalytic`

Simple, standard deviation based normalcy calculation. Normalcy can be based on 1 or more attributes in an event. Plugin is fully concurrent, meant for high throughput (50k+ EPS) installs.

Property	Description	Data Type or Valid Values
attribute_xpath.x	XPaths to identifying attributes	XPath string

<filter>	GeoAnalyticFilter	
aggregate_classifications	Whether to accumulate and combine classifications and data groups	boolean
history_file_name	File to store history	File path
training_time_s	Waiting period for the analytic to learn normal levels before normalcy analytic events are produced.	Postive long or hh:mm:ss string
window_size_s	Rolling window size for normalcy calculation	Postive long or hh:mm:ss string
max_data_lag_s	Maximum delay in data arrival from event time. The analytic will wait this amount of time for data to come in before calculating normalcy on the period.	Postive long or hh:mm:ss string
normalcy_threshold	Maximum normalcy value to sent analytic events on, between 0 (very abnormal) and 1 (completely normal), inclusive.	Double 0-1
history_size_periods	The number of historical periods to use for normalcy calculation	Positive int

Table 36 - Simple Normalcy Analytic options

Output Property	Description	Data Type
id	ID of abnormal attribute	String
value	Observed value	double
normalcy	0 to 1 normalcy of value	double
historicalValues	Comma separated list of historical values	Comma separated list of doubles
mean	Average value	double
numDeviations	Number of deviations the value is from the mean	double
stdDev	The standard deviation of observed values	double
startDate	Start date of observation period	Date string
endDate	End date of observation period	Date string

Simple Normalcy Analytic output properties

2.1.34 Rate of Growth Normalcy Analytic

`icg.engine.analytic.rog.normalcy.RateOfGrowthNormalcyAnalytic`

Measures the rate of growth of entities from first being registered with the system. Produces a periodic report of the entities with the top growth rate in both gross volume and maximum percentage spike.

Property	Description	Data Type or Valid Values
event_stream	The stream name that events to count are on	String
registration_stream	The stream name that registration events are on	String
num_results	The number of top results to report	int
num_extra_buckets_before_expire	The number of time buckets to keep completed entries in the comparison population	int
num_buckets	The number of time buckets before an entry is measured	int
num_skip_buckets	The number of initial time buckets to skip in a comparison	int
max_data_lag_s	The maximum amount of seconds that events are lagged from real time	Postive long or hh:mm:ss string
bucket_size_s	The size of a bucket in seconds	Postive long or hh:mm:ss string
aggregate_classifications		Boolean
event_id_xpath	Xpath to the entity ID in the event stream	XPath String
registration_id_xpath	XPath to the entity ID in the registration stream	XPath String
.registration	Postfix for registration stream's AnalyticFilter's properties	
.event	Postfix for event stream's AnalyticFilter's properties	

Rate of Growth Normalcy Analytic options

Output Property	Description	Data Type
pop_size	Size of observed population	double
mean_count	Average count of population	double
mean_growth_rate	Average growth rate of population	double
top_growth_rate_x_id	ID of item x in top growth rate chart	String
top_growth_rate_x	Growth rate for item x in top growth rate chart	double
top_count_x_id	ID of item x in top count chart	String
top_count_x	Count for item x in top count chart	double

Rate of Growth Normalcy Analytic output properties

2.1.35 Rendezvous Analytic (Geohash Clustering based)

`icg.engine.analytic.geohashclustering.RendezvousAnalytic`

Alerts when it detects a rendezvous. The data is grouped by whatever element is specified in configuration. Rendezvous must contain a minimum configurable number of unique ids, must occur within a specified time window, must be bounded by a circle of no more than a specified size, and must contain posits traveling at approximately the same speed and in approximately the same direction.

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
course_xpath	XPath to course 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_course_deviation_percent	Max deviation from first course in cluster to consider event	Positive real [0-100]
max_speed_deviation_percent	Max deviation from first speed in cluster to consider event	Positive real [0-100]
max_events	Max number of events in a cluster	positive integer
maximum_speed	Maximum value of speed attribute to consider the event	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
min_ids	Minimum number of unique ids in cluster.	Positive integer
minimum_speed	Minimum value of speed attribute to consider the event	Positive integer
speed _xpath	XPATH of speed attribute	XPath string

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 from first event to current event	Positive long or hh:mm:ss string
<filter>	GeoAnalyticFilter	
cluster_style_name	Style name for cluster bounding circle, default `cluster`	String

Rendezvous Analytic options

Output Property	Description	Data Type
cluster_area	Geo bounds of cluster	GeoJSON circle
clusterCenter	Center geo for cluster	GeoJSON point
clusterCenterLatitude	Latitude of center of cluster	double
clusterCenterLongitude	Longitude of center of cluster	double
diameter	Diameter of cluster in meters	double
totalDistanceMeters	Total distance between all events in the cluster, in meters	double
clusterStartTime	Time of earliest event in the cluster	long
clusterDurationMs	Time delta between earliest and latests cluster events, in milliseconds	long
common_attribute	The common attribute value for events in the cluster, if set	String
description	Description string including common attribute	String
number_unique_ids	Number of unique ids in the cluster, if configured	int
unique_ids	Each unique ID in the cluster	String

Rendezvous Analytic output properties

2.1.36 Inactive Forecast Analytic

`icg.engine.analytic.track.forecast.InactiveForecastAnalytic`

Creates forecasts on tracks that have gone inactive, attempting to predict where the entities are at the time of prediction. Uses `GeoGridTrackForecastAnalytic` logic to make prediction if possible, else falls back on a dead reckoning via course and speed attributes.

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	5003500, 625400, 123260, 19540, 3800, 605, 116, 18, 3, 0.5

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. Must be greater than or equal to min_geohash_track_size_for_prediction	int
min_geohash_track_size_for_prediction	Minimum number of geohashes present in a track for a prediction to be made, default 2. Must be less than or equal to max_geohash_chain_size.	int
<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	Postive long or hh:mm:ss string
track_style_name	Style name for track line, default `track`	String
track_point_style_name	Style name for track point, default `track_point`	String
prediction_area_style_name	Style name for prediction square, default `GeoGridTrackForecast`	String
prediction_path_style_name	Style name for line between prediction squares, default `predicted_path`	String
history_file_name	Optional, file to save learned history to	File path string
report_incorrect_predictions	If true, the analytic will produce events for entities that do not follow their forecasted track. Default false.	Boolean
save_interval_s	Number of seconds between history saves. Optional. Default: `00:10:00`	Postive long or hh:mm:ss string
min_inactive_time_s	Minimum time an entity can be inactive and still generate forecasts, in seconds	Postive long or hh:mm:ss string
max_inactive_time_s	Maximum time an entity can be inactive and still generate forecasts, in seconds	Postive long or hh:mm:ss string

inactive_check_interval_s	Frequency to check for inactive entities and produce forecasts, in seconds	Postive long or hh:mm:ss string
process_out_of_order_updates	If false, old events (out of order) will be ignored. Default `false`	Boolean
dr_point_style_name	Style name for dead reckoning points, default `dr_point`	String

Inactive Forecast Analytic options

Output Property	Description	Data Type
description	Description string including ID	String
calc_speed_ms	Calculated average speed in m/s	double
track_length_m	Length of the track in meters	double
track_duration_hhmmss	Duration of the track	Hh:mm:ss string
track_point_x	Geo for point x on track	GeoJSON point
track	Line for track	GeoJSON line
predictionPath	Line for prediction	GeoJSON line
predictionArea	Predicted area (can be more than one)	GeoJSON polygon

Inactive Forecast Analytic output properties

2.1.37 Entity State Machine Analytic

`icg.engine.analytic.esm.EntityStateMachineAnalytic`

Allows definition of a chain of n states for an entity to transition through before an analytic event is generated. States are defined by a time period and n event attribute conditions that must be satisfied. Once a state is satisfied, the entity moves to the next state. If the final state is satisfied, an analytic event is created. If all states aren't satisfied in `total.duration.max.s` seconds, state is reset to the initial state.

Property	Description	Data Type or Valid Values
location_xpath	XPath to location	XPath string
id_attribute_xpath	XPath to ID attribute	XPath string
id_attribute_output_property	Property name for ID attribute	String
id_attribute_display_name	Display name for ID attribute	String
description_format_string	Single string variable format string. Default: `%s has met all conditions for + getDisplayName()`	Format string
write_to_graph_store	Whether or not to write entities and relationships to the Entity Manager Graph Store	Boolean

total_duration_max_s	Total duration for all states to be satisfied before entity state is reset	Postive long or hh:mm:ss string
state.x.attribute.y.**xpath**	XPath to attribute y in state x	XPath string
state.x.attribute.y.**op**	Operation for attribute y in state x. `delta_max` - specifies that observed values of this attribute must stay within a specified delta of each other for the duration of the state `range` - specifies a min and max value that the attribute must stay within for the duration of the state	`delta_max` or `range`
state.x.attribute.y.**arg**	`delta_max` - double range e.g. `10` `range` - min/max (inclusive) double values, comma-separated. E.g. `4,8`	String, depending on operation
state.x.duration_s	The amount of time in seconds that a state's conditions must be satisfied before it moves on to state x+1, or sends an analytic event if this is the last state	Postive long or hh:mm:ss string
<EntityConfig>	If write_to_graph_store is true	

Entity State Machine Analytic options

Output Property	Description	Data Type
description	Description string including ID	String
id	ID of entity that matched criteria	String
entity_id	If entity manager is enabled, ID of entity from entity database	String

Entity State Machine Analytic output properties

Example configuration:

```
location_xpath=/event/georss:where
id_attribute_xpath=/event/id
state.1.attribute.1.xpath=/event/speed
state.1.attribute.1.op=range
state.1.attribute.1.arg=10,20
state.1.attribute.2.xpath=/event/course
state.1.attribute.2.op=delta_max
state.1.attribute.2.arg=10
state.1.duration_s=01:00:00
state.2.attribute.1.xpath=/event/speed
state.2.attribute.1.op=range
state.2.attribute.1.arg=4,8
state.2.attribute.2.xpath=/event/course
state.2.attribute.2.op=delta_max
state.2.attribute.2.arg=10
```

`state.2.duration_s=01:00:00`
`total_duration_max_s=24:00:00`

Output Property	Description	Data Type
analytic_entity_state_machine_time	Time of analytic event	DATESTRING
analytic_entity_state_machine_location	Location of analytic event	GEO
analytic_entity_state_machine_description	Display name of analytic	String

Entity State Machine Analytic EM Entity properties

2.1.38 Track Gap Proximity Analytic

`icg.engine.analytic.track.gap.TrackGapProximityAnalytic`

Looks for gaps in tracks of a minimum time delta. Once found, uses a* pathfinding to interpolate between points. Uses interpolated tracks to look for proximities that may have occurred, and alerts on any found. TRACE level logging will write the interpolated tracks to disk in KML format.

Property	Description	Data Type or Valid Values
location_xpath	XPath to location	XPath string
id_attribute_xpath	XPath to ID attribute	XPath string
max_track_length	Maximum number of posits per track to hold in the track manager component	int
max_point_age_s	Max age in seconds of points to keep for analysis	long
min_gap_time_s	Minimum track gap in seconds needed to interpolate and look for proximities	long
<filter>	GeoAnalyticFilter	
aggregate_classifications	Whether to accumulate and combine classifications and data groups	boolean
proximity_distance_threshold_meters	Maximum distance needed between entities for an analytic event to be created	double
proximity_time_threshold_seconds	Maximum amount of time between projected track posit and actual posit for a proximity event to be created	long

geohash_data_file	File containing geohash data for obstacle avoidance. Data file format is csv: geohash,value	File path
a_star_geohash_chars	Number of geohash characters to use in pathfinding coordinates. More characters = more squares = longer processing times. Default is the highest resolution in the `geohash_data_file` + 1.	int
LAND	"Land" is the obstacle value used in the `geohash_data_file`'s value column	
a_star_geohash_divisor	How many lat/lon slices to create as options for pathfinding within each geohash square. Higher divisor = more points = longer processing times. Default is 3	int
prediction_track_style_name	Style name for interpolated track	String
point_style_name	Style name for proximity points	String
write_to_graph_store	Whether or not to write entities and relationships to the Entity Manager Graph Store	Boolean
<RelationshipConfig properties>	If write_to_graph_store is true	

Track Gap Proximity Analytic options

Output Property	Description	Data Type
description	Description string including ID	String
gap_id	ID of the entity whos track was interpolated	String
prox_id	ID of the entity in proximity of `gap_id` entity	String
gap_id_point	Location of `gap_id` entity when the proximity occurred	GeoJSON point
prox_id_point	Location of `prox_id` entity when the proximity occurred	GeoJSON point
gap_event_time	Event time of `gap_id` entity event	Date string
prox_event_time	Event time of `prox_id` entity event	Date string
relationship_id	If entity manager is enabled, ID of relationship from entity database	String

Track Gap Proximity Analytic output properties

Output Property	Description	Data Type
<Relationship>	Relationship between track gap entity	

	and proximity entity	

Track Grap Analytic EM properties

2.1.39 Indicators Analytic

`icg.engine.analytic.indicators.IndicatorsAnalytic`

Combines probability from a number of configured independent indicator events. Each indicator is an attribute/value match on a given stream, and is assigned a probability. The probability can change over time as configured. All probabilities are combined into an overall score, representing the likelyhood of an event occuring.

Property	Description	Data Type or Valid Values
min_report_time_seconds	Minimum time before a score report will be generated. It will also be generated when an event matches an indicator.	Postive long or hh:mm:ss string
indicator_x_id	ID for indicator x	String
indicator_x_match_value	Attribute in indicator_x_xpath must match this value for the indicator to activate	String
indicator_x_xpath	Path to match attribute	XPath string
indicator_x_stream	Stream for indicator_x to monitor	String
<filter>	GeoAnalyticFilter	
aggregate_classifications	Whether to accumulate and combine classifications and data groups	boolean
indicator_x_score_function_type	Score modification function. Scores can increase or decrease with time after in indicator is activated	`step` or `linear`
indicator_x_score_function_initial_score	For `linear` score function, initial score after activation	Double 0-1
indicator_x_score_function_delta_per_second	For `linear` score function, score delta per second after activation	Double 0-1
indicator_x_score_function_zero_threshold	For `linear` score function, the score at which the value is considered 0.	Double 0-1
indicator_x_score_function_step_y_time	For `step` score function, time after activation for step y	long
indicator_x_score_function_step_y_value	For `step` score function, value for step y	Double 0-1

Indicators Analytic options

Output Property	Description	Data Type

overall_score	Combined probability score for this report	Double 0-1
report_time	Time this report was generated	Date long
indicator_id_x	See Indicators Analytic options table	
indicator_score_x	Current score for indicator x	Double 0-1
indicator_obs_time_x	Time indicator x was last activated	Date long
indicator_score_function_x	See Indicators Analytic options table	
Indicator_score_function_x_initial_score	See Indicators Analytic options table	
indicator_score_function_x_delta_per_second	See Indicators Analytic options table	
indicator_score_function_x_zero_threshold	See Indicators Analytic options table	
indicator_score_function_x_range_y_lower	Lower bound for step y	long
indicator_score_function_x_range_y_upper	Upper bound for step y	long
indicator_score_function_x_range_y_value	Value for step y	Double 0-1

Indicators Analytic output properties

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	boolean

	filtered text events for age	

Table 29 - Age Prediction Enrichment options

Output Property	Description	Data Type
<output_property>	The determined age	int
age_bracket	The determined age range	<int>-<int>

Age Prediction Enrichment output properties

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 30 - Gender Prediction Enrichment options

Output Property	Description	Data Type
<output_property>	Determined gender of author	String

Gender Prediction Enrichment output properties

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, use this or kml_area_file	File path
data_file_poll_interval_s	How frequently to poll the data file for changes, in seconds. Default is	Postive long or hh:mm:ss string

	no polling.	
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_xpath	XPath to event location	XPath string
hit_distance_property	Property name for distance to ID	string
enable_asset_geos	If true, will output the matching geo IDs and shapes with events. Default: true.	boolean
kml_areas_file	KML file for shapes, use this or data_file	File path

Table 31 - Geo Proximity Enrichment options

Output Property	Description	Data Type
entity_hit.<hit_output_property>	Hit ID	String
entity_hit.<hit_distance_property>	Event distance to ID in meters	double
hit_circle	Circle around ID with radius equal to its configured threshold	Geo
hit_center	Hit ID location	Geo

Geo Proximity Enrichment output properties

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_xpath	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	Postive long or hh:mm:ss string
hit_output_property	Property name for a location hit	string
hit_output_property_display_name	Display name for a location hit	string

Table 32 - Geo Tagging Enrichment options

Output Property	Description	Data Type
<output_property>	ID of hit	String

Xyz Enrichment output properties

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 33 - Group Membership Enrichment options

Output Property	Description	Data Type
<group_output_property>	Group with the most connections	String
<connection_output_property>	The number of connections to the group	int

Group Membership Enrichment output properties

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, .1, .01, .001, etc

	.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_s	Optional - frequency with which to save histogram state, in seconds. Default - `00:05:00`	Positive long or hh:mm:ss string

Table 34 - HBOS Anomaly Enrichment options

Output Property	Description	Data Type
hbos_anomaly_score	HBOS score	double

HBOS Anomaly Enrichment output properties

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 35 - IP Address Geoservice Enrichment options

Output Property	Description	Data Type
<output_property>	Geo lookup results	String

Xyz Enrichment output properties

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 36 - IP to Geo Enrichment options

Output Property	Description	Data Type
geo	Geo point	Geo
city	The city	String
state	The state	String
country	The country	String

IP to Geo Enrichment output properties

2.2.8 Fast IP to Geo Enrichment

`icg.engine.enrichment.iptogeo.FastIPToGeoEnrichment`

Fully concurrent IP-to-geo lookup meant for high volume (10k+ EPS) systems. Requires a **`fastiptogeo.conf`** file in the data directory with a single line, an xpath query to the IP address field.

Property	Description	Data Type or Valid Values
None (see description)		

Table 37 - Fast IP to Geo Enrichment options

Output Property	Description	Data Type
geo	Geo point	Geo
city	The city	String
state	The state	String
country	The country	String
country_code	The two-character ISO 3166-1 alpha code for the country	String

IP to Geo Enrichment output properties

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
write_to_graph_store	Optional, if true, detected language can be added as an entity attribute	boolean

output_property	Output property for detected language, default **`language`**	string
output_property_display_name	Display name for detected language property, default **`Language`**	string

Table 36 - Language Detection Enrichment options

Output Property	Description	Data Type
<output_property>	Detected language	String

Language Detection Enrichment output properties

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

Output Property	Description	Data Type
last_observation	Time since last observation, in secontds. Or “N/A”	String

Last Observation Enrichment output properties

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

Output Property	Description	Data Type
<title_output_property>	Article title	String
<text_output_property>	Article text	String

Link Fetcher Enrichment output properties

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

Output Property	Description	Data Type
<output_property>	Output property as configured	String

Matching Engine Enrichment output properties

2.2.13 NLP Enrichment

`icg.engine.enrichment.nlp.NLPEnrichment`

Uses the University of Illinois NER library conduct 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
output_prefix	Optional prefix for output properties, default is "`nlp.`"	String

Table 40 - NLP Enrichment options

Output Property	Description	Data Type
<output_prefix>MAIN_PERSON	The main person	String
<output_prefix>MAIN_LOCATION	The main location	String
<output_prefix>MAIN_LOCATION.geo	The geo center of the main location	Geo
<Output from U Ill lib>		

NLP Enrichment output properties

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

Output Property	Description	Data Type
<output_property>	Determined political party	String

Political Party Enrichment output properties

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

Output Property	Description	Data Type
<output_property>	Name or number of capturing group	String

Regex Capturing Group Enrichment output properties

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

Output Property	Description	Data Type
<output_property>	The modified text	String

Regex Replacement Enrichment output properties

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

Output Property	Description	Data Type
regex_match	Matching text	String

Regex Substring Enrichment output properties

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

Output Property	Description	Data Type
<output_property>	Output property as configured	String

Shared Data Enrichment output properties

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

Output Property	Description	Data Type
<output_property>	One of the split tokens	String

Splitter Enrichment output properties

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 beginning of the word, that must match to consider two words equal.	integer
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

Output Property	Description	Data Type
<output_property>	Stock symbols found	String

Stock Symbol Enrichment output properties

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

Output Property	Description	Data Type
<output_properties from csv file>		

Translation Service Enrichment output properties

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

Output Property	Description	Data Type
<reply_screenname_property>		String
<mentioned_property>		String
<hashtag_property>		String
<cashtag_property>		String
<url_property>		URL
profile_age_hours	Age the profile has existed, in hours	double
followers_per_hour	Number of followers gained per hour since the profile was created	double
following_per_hour	Number of profiles this user has followed per hour since it was created	double

Xyz Enrichment output properties

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
output_property	Output property for influence score, default `influenceScore`	String
output_property_display_name	Display name for influence score output property, default `Author Influence Score`	String
write_to_graph_store	If true, will write influence score to	Boolean

	graph store Entities	

Table 50 - Twitter Influence Enrichment options

Output Property	Description	Data Type
<output_property>	Influence score	double
avg_retweets	Average number of retweets for this user's tweets	double
avg_tweet_interval	Average tweet interval	HH:MM:SS string

Twitter Influence Enrichment output properties

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

Output Property	Description	Data Type
main_subject	The main subject	String

Twitter Main Subject Enrichment output properties

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
write_to_graph_store	If true, will write influence score to graph store Entities	Boolean

Table 52 - UofI Ethnicity Enrichment options

Output Property	Description	Data Type
<main_ethnicity>	Ethnicity	String

UofI Ethnicity Enrichment output properties

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

Output Property	Description	Data Type
<main_sentiment_output_property>	Detected sentiment	String

UofI Sentiment Enrichment output properties

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

Output Property	Description	Data Type
<expanded_urls_property>	Expanded URLs if configured	URL

site_name	Domain name, if bias configured	URL
site_rating	Bias rating from file if configured	String

Xyz Enrichment output properties

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

Output Property	Description	Data Type
url_host		String
url_port		int
url_path		String
url_protocol		String
url_query		String

URL Splitter Enrichment output properties

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
value_format	Method to parse mapped values (including default_value, static values, and value column contents from csv)	TEXT (default), NUMBER, LAT_LON, LON_LAT

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_s	How frequently to poll the CSV file for updates, in second. Default `00:05:00`	Positive long or hh:mm:ss string
value_path	XPath to value attribute in event	XPath string
match-group.csv.header.lines	Optional - number of header lines in csv file	int
match-group.csv.key	Optional - static value for mapping key	String
match-group.csv.key.column	Optional - column to find the mapping key	int
match-group.csv.value	Optional - static value for mapping value	String
match-group.csv.value.column	Optional - column to find the mapping value	int

Table 56 - Value Map Enrichment options

Output Property	Description	Data Type
<output_property>	The mapped value	String

Value Map Enrichment output properties

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

Output Property	Description	Data Type

<output_property>	The value range	String

Xyz Enrichment output properties

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

Output Property	Description	Data Type
<output_property>	Detected sentiment	String
<emotion_output_property>	Detected emotion if configured	String

VS Sentiment Enrichment output properties

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

Output Property	Description	Data Type
subject_x_similarity	Similarity to subject x	String

Word2Vec Enrichment output properties

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

Output Property	Description	Data Type
<output_property>	OCR text	String

OCR Enrichment output properties

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 course attribute in degrees	XPath string
speed_knots_xpath	XPath to speed attribute in knots	XPath string
location_xpath	XPath to location attribute	XPath string
prediction_seconds_x	Number of seconds to predict	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

Output Property	Description	Data Type
geo_<time_s>	Geo for time_s time	Geo

Course Speed Projection Enrichment output properties

2.2.34 Event Age Enrichment

`icg.engine.enrichment.event.age.EventAgeEnrichment`

Adds the age of the event in the selected units, optionally rounded. Default is hours, not rounded.

Property	Description	Data Type or Valid Values
time_unit	Time unit for the output property	`DAYS`, `HOURS`, `MINUTES`, `SECONDS`, or `MILLISECONDS`
round_down_enabled	If true, results will be rounded down to the nearest integer, default `false`	Boolean
<filter>	GeoAnalyticFilter	

Table 62 - Event Age Enrichment options

Output Property	Description	Data Type
event_age	Age of event in chosen time units	double

Event Age Enrichment output properties

2.2.34 Substring Enrichment

`icg.engine.enrichment.substring.SubstringEnrichment`

Parses the first n characters, words, sentences, or paragraphs from a text attribute and makes them their own property.

Property	Description	Data Type or Valid Values
number	The number of text units to parse	Positive integer
text_unit	The unit of text to parse	`CHARACTER`, `WORD`, `SENTENCE`, or `PARAGRAPH`
text_attribute_path	Path to text attribute	XPath string
output_property_name	Name for output property	String
output_property_display_name	Display name for output property	String

Table 63 - Substring Enrichment options

Output Property	Description	Data Type
<output_property>	The substring	String

Substring Enrichment output properties

2.2.34 Domain Registration Enrichment

`icg.engine.enrichment.domainregistration.DomainRegistrationEnrichment`

Takes a stream of registration events and a stream of network events. Saves host/domain associations from registration stream. Adds hosts to network event stream based on domain.

Property	Description	Data Type or Valid Values
registration_stream	Name of stream that registration events are on	String
event_stream	Name of stream that events are on	String
domain_path	Path to domain in reg and event streams	XPath string
host_path	Path to host attribute in reg stream	XPath string
output_property_name	Name for output property	String
output_property_display_name	Display name for output property	String
<event filter> `.registration`	Event filter properties for registration stream	
<event filter> `.event`	Event filter properties for event stream	

Domain Registration Enrichment options

Output Property	Description	Data Type
host		String
registrar		String
registrant_name		String
regisrant_country		String
created_date		String

Domain Registration Enrichment output properties

2.2.34 Multi-Stream Value Map Enrichment

`icg.engine.enrichment.valuemap.MultiStreamValueMapEnrichment`

Adds a properties/attributes from source stream to events in the dest stream based on a common identifier.

Property	Description	Data Type or Valid Values
source_stream	Name of stream that attributes will be read from	String
dest_stream	Name of stream that properties will be added to	String
source_id_path	Path to id in source stream	XPath string
dest_id_path	Path to id in dest stream	XPath string
source_att_query.x	Path to attribute x in source stream	XPath string
output_property.x	Output property for attribute x	String
output_property_display_name.x	Output property display name for attribute x	String

<event filter> `.source`	Event filter properties for source stream	
<event filter> `.dest`	Event filter properties for dest stream	
aggregate_classifications	Whether to combine the classifications of all events in all paths for the output classification	boolean
history_file_name	Optional, file to save mappings to	File path string
save_interval_s	Optional, frequency in seconds to save mappings to history file, default `300`	Positive long or hh:mm:ss string

Multi-Stream Value Map Enrichment options

Output Property	Description	Data Type
<output_property>	The mapped value	String

Multi-Stream Value Map Enrichment output properties

2.2.35 Closest Proximity Enrichment

`icg.engine.enrichment.closestproximity.ClosestProximityEnrichment`

Loads points of interest from a CSV file. Enriches events with the closest point of interest to the event, if any are less than `distance_threshold_meters`. CSV file must have rows for ID, lat, and lon. Optionally, the enrichment will add additional properties from the CSV file to the event.

Property	Description	Data Type or Valid Values
location_xpath	XPath to location attribute	XPath string
data_file_poll_interval_s	Frequency in seconds to poll the CSV file, default `300`	long
distance_threshold_meters	Distance	long
hit_id_output_property	Property for ID of closest point	String
hit_id_output_property_display_name	Display name for ID of closest point	String
hit_distance_output_property	Property for distance to closest point	String
hit_distance_output_property_display_name	Display name for distance to closest point	
<filter>	GeoAnalyticFilter	
id_index	CSV column index of point ID	int
lon_index	CSV column index of longitude	int
lat_index	CSV column index of latitude	int
attribute.x.name	Optional, property name for attribute x from CSV file	String
attribute.x.display_name	Optional, display name for attribute x from CSV file	String
attribute.x.index	Optional, column index for attribute x	int

num_header_lines	Optional, number of CSV lines to skip, default 0	int
data_file	Path to CSV file	File path string

Closest Proximity Enrichment options

Output Property	Description	Data Type
entity_hit.<hit_output_property>	Hit ID	String
entity_hit.<hit_distance_property>	Event distance to ID in meters	double
hit_circle	Circle around ID with radius equal to its configured threshold	Geo
hit_center	Hit ID location	Geo

Closest Proximity Enrichment output properties

2.2.36 Track ID Enrichment

`icg.engine.enrichment.trackid.TrackIDEnrichment`

Assigns unique track IDs based on a semi-unique identifier, such as MMSI. Attempts to resolve duplicate IDs via geospatial calculations based on the **`max_speed_meters_per_second`** property.

Property	Description	Data Type or Valid Values
location_xpath	XPath to location attribute	XPath string
id_attribute_xpath	XPath to semi-unique identifier attribute	XPath string
output_property	Optional, Output property, default **`track_id`**	String
output_property_display_name	Optional, Display name for output property, default **`Track ID`**	String
history_file_name	Optional, file path to save history	File path string
save_interval_s	Optional, number of seconds between history saves, default **`600`**	Long or hh:mm:ss
max_speed_meters_per_second	Max speed in meters per second that entities of this type can travel	double
<filter>	GeoAnalyticFilter	
geo_error_m	Optional, if specified, distances of less than this amount will be ignored when de-duping. This will prevent sensor error from splitting tracks.	Double

Track ID Enrichment options

Output Property	Description	Data Type
<output_property>	Track ID	String

Track ID Enrichment output properties

2.2.37 Rekognition Enrichment

`icg.engine.enrichment.rekognition.RekognitionEnrichment`

Labels photos containing celebrities (qtype celebrity) or objects (qtype label) using Amazon Rekognition API.

Property	Description	Data Type or Valid Values
poll.timeout	Value used in http request for timeout, in milliseconds, default 3000	Positive integer
amazon_user	Identifies credentials, stored in ~/.aws/credentials file under [<amazon_user>] tag. (/usr/local/lux/.aws)	String
max_image_size_byes	Optional, default 3MB. Maximum size of the image file in bytes. Images larger than this will be resized to contain no more than this. Amazon's maximum image size is 5MB.	Positive integer
qtype_	Query Type prefix, followed by a sequential integer (e.g., qtype_1) Query Type value can be one of the following: label, celebrity, dynamic. (See Amazon Rekognition API).	String
dynamic_qtype_path_	Dynamic QType Path Prefix, followed by a sequential integer. If qtype is dynamic, this is the XPath to the field that contains one of the following strings: #label#, #celebrity#. If neither string is found, qtype is both.	XPath String
query_	XPath prefix (followed by a sequential integer) to the image (starts with data:image/png;base64) or image URL (starts with http)	XPath String
min_confidence_	Minimum confidence prefix (followed by a sequential integer). Minimum confidence required to emit label property in output	String

Rekognition Enrichment options

Output Property	Description	Data Type
output_property_	Output Property Prefix (followed by a sequential integer). Name to use for all of the values returned by rekognition.	String
output_property_display_name_	Output Property Display Name Prefix (followed by a sequential integer). Display name for the name property	String

	above.	

Rekognition Enrichment output properties

2.2.38 LDA Enrichment

`icg.engine.enrichment.lda.LDAEnrichment`

Uses the Latent Dirichlet Allocation (LDA) algorithm to extract topics and keywords from text fields.

Property	Description	Data Type or Valid Values
text_path	XPath to text attribute	XPath string
num_keywords	The number of top keywords to output in **`keyword`** and **`top_keywords`** properties	integer
<filter>	GeoAnalyticFilter	

LDA Enrichment options

Output Property	Description	Data Type
topic_summary	Summary of topic	String
keyword	A keyword	String
top_keywords	CSV of top keywords	CSV String

LDA Enrichment output properties

2.2.39 Topic Enrichment

`icg.engine.enrichment.lda.TopicEnrichment`

Uses the Latent Dirichlet Allocation (LDA) algorithm to extract keywords from text fields, then uses FastText word vectors to score keywords against predefined topic keywords from config file.

Property	Description	Data Type or Valid Values
text_path	XPath to text attribute	XPath string
num_keywords	The number of top keywords to output in **`keyword`** and **`top_keywords`** properties	integer
<filter>	GeoAnalyticFilter	
num_matches	The number of topic matches to add, default 3	integer
topic.x.name	The name for topic x	String
topic.x.keywords	Comma separated list of keywords for topic x	Comma separated list
fasttext_model_path	Path to fasttext bin model	File path

Topic Enrichment options

Output Property	Description	Data Type
topic_match_<topic_name>	Percent match to topic	double
keywords	Keywords for top topic match	CSV String

Topic Enrichment output properties

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 64 - 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

lux.alert.format.title	Title format string for alert	Alert title format string
lux.alert.json.template.dir	Optional, path to folder containing FTL template	File path
lux.alert.json.template	Optional, FTL template file name to use to convert alerts. Omit ".ftl" from file name.	File name string
alerter.param.to.add.x	Keys for alerter parameters to add to alerts as properties.	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 JSON 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	Optional, 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	string

	parameter can be ignored.	
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
event.parser	Class name of LUXEventParser to use	Fully qualified class name
event.ingest.plugin.path	Optional, file path to location of jars to add to classpath	File path

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

event.parser	Class name of LUXEventParser to use	Fully qualified class name
event.ingest.plugin.path	Optional, file path to location of jars to add to classpath	File path

Table 74 - Email Ingest options

Property	Description	Data Type or Valid Values
email_event_class_name	Fully qualified name of event class	String

icg.engine.event.ingest.data.parsers.email.LuxEmailParser properties

Property	Description	Data Type or Valid Values
encoding	Optional, Name of encoding	Default UTF-8
use_event_time	Optional, use event time	Default true
max_image_size_bytes	Optional, Maximum size of image in bytes to include in event from an attached or inline image.	Positive integer, default 2,097,152 bytes (2MB)

icg.engine.event.ingest.data.parsers.email.LuxEmailEvent properties

Property	Description	Data Type or Valid Values
encoding	Optional, Name of encoding	Default UTF-8
use_event_time	Optional, use event time	Default true

icg.engine.event.ingest.data.parsers.email.PiracyEmailEvent properties

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,	boolean

	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	
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
event.parser	Class name of LUXEventParser to use	Fully qualified class name
event.ingest.plugin.path	Optional, file path to location of jars to add to classpath	File path
file.preprocessor	Optional, FilePreprocessors get the files first and convert them in some way, then return a new File pointer.	Class name

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`
event.parser	Class name of LUXEventParser to use	Fully qualified class name
event.ingest.plugin.path	Optional, file path to location of jars to add to classpath	File path

Table 77 - FTP File Ingest options

2.4.7 JMS Ingest

`icg.engine.jms.JMSIngest`

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. Specify either jms.queue.name or jms.topic.name.	string
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
event.parser	Class name of LUXEventParser to use	Fully qualified class name
event.ingest.plugin.path	Optional, file path to location of jars to add to classpath	File path

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
event.parser	Class name of LUXEventParser to use	Fully qualified class name
event.ingest.plugin.path	Optional, file path to location of jars to add to classpath	File path
event.parser	Class name of LUXEventParser to use	Fully qualified class name
event.ingest.plugin.path	Optional, file path to location of jars to add to classpath	File path

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. Supports SSL.

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	boolean

	new data. Default false	
topic	Kafka topic to read from	string
zookeeper.connect	Zookeeper connect URL, only needed if `start.from.beginning` is `true`	URL
ssl.enabled	Flag for enabling SSL, if `true`, all `ssl.*` properties are required	Boolean
ssl.keystore.path	File path to keystore	File path
ssl.truststore.path	File path to truststore	File path
ssl.keystore.password	Password for keystore	String
ssl.key.password	Password for key	String
ssl.truststore.password	Password for truststore	String
event.parser	Class name of LUXEventParser to use	Fully qualified class name
event.ingest.plugin.path	Optional, file path to location of jars to add to classpath	File path

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
event.parser	Class name of LUXEventParser to use	Fully qualified class name
event.ingest.plugin.path	Optional, file path to location of jars to	File path

	add to classpath	

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`
event.parser	Class name of LUXEventParser to use	Fully qualified class name
event.ingest.plugin.path	Optional, file path to location of jars to add to classpath	File path

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	JSON string

	parser will drill down recursively until it finds arrays of numbers.	
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 Event Parser

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

Parses CSV rows from an InputStream.

Property	Description	Data Type or Valid Values
event.id.name	Csv column header name of the event id. If not specified, the event id is generated.	String
event.date.name	Csv column header name of the event time. If not specified, the event time is the current time.	String
header.names	Csv column header names, defines the format of the csv input	Comma separated list of Strings
active.header.names	Csv column header names of the active columns - the columns to add to the event	Comma separated list of Strings
integer.names	Csv column header names of the fields that should be interpreted as of type Integer.	Comma separated list of Strings
double.names	Csv column header names of the fields that should be interpreted as of type Double.	Comma separated list of Strings.
date.names	Csv column header names of the fields that should be interpreted as of type Date.	Comma separated list of Strings.
rights.name/rights.value	Csv column header name of the rights or the (constant) value of the rights. One or the other but not both must be specified.	String
data.groups.value	The constant value of the data groups.	Comma separated list of Strings.
description.value	The constant value of the description.	String
publisher.name/publisher.value	Csv column header name of the event publisher or the (constant) value of the publisher. If not specified, publisher is null.	String
source.name/source.value	Csv column header name of the event source or the (constant) value of the source. If not specified, source is null.	String
type.name/type.value	Csv column header name of the event type or the (constant) value of the type. If not specified, type is null.	String

title.name/title.value	Csv column header name of the event title or the (constant) value of the title. If not specified, title is null.	String
geom.name/lat.name/lon.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.	String
supported.date.format	prefix (add a 1, 2, 3, ...) for the supported date format string, or if none, uses defaults	Java DateFormat strings. Defaults are "yyyy-MM-dd'T'HH:mm:ss.SSS'Z'", "yyyyMMdd'T'HH:mm:ss.SSS'Z'", "yyyyMMddHHmm"};
delim	Optional non-comma delim character	Character

Table 86 - CSV File Parser options

2.5.2 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 87 - CSV File Parser options

2.5.3 Generic JSON Parser

`icg.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 the event publisher. Specify publisher.name or publisher.value.	Dot notation path
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	Dot notation path

	parsed. All required attributes must be present.	
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
make.lla.props	If true, lat/lon/alt will be parsed from geometry object and added as event attributes	Boolean

Table 88 - Generic JSON Parser options

2.5.4 Simple LUX Event Parser

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

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

2.5.5 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 89 - Streaming HTML Parser options

2.5.6 TACREP Parser

`icg.engine.event.ingest.data.parsers.usmtf.TACREPParser`

Parses TACREPs in / format. Expecting either one event per input stream or one event per line, toggled with `one.event.per.stream` flag.

Property	Description	Data Type or Valid Values
one.event.per.stream	Whether each stream contains a single event, or one event per line	Boolean

TACREP Parser options

2.5.7 NMEA/NM4 AIS Event Parser

`icg.engine.event.ingest.data.parsers.nmea.NM4EventParser`

Parses NMEA / NM4 / OTHG AIS messages delivered in CSV format, with each message contained in a single row. Supports the following message types: 1-5, 9, 11, 12, 14, 18, 19, 21, 24, and 27. Message type reference https://www.navcen.uscg.gov/?pageName=AISMessages

2.5.8 OTHG Event Parser

`icg.engine.event.ingest.data.parsers.othg.OTHGEventParser`

Parses OTHG events in forward-slash delimited format.

Property	Description	Data Type or Valid Values
use_event_time	If true, event time will be the time in the event rather than current time. Default **`false`**	boolean
do_partial_line_processing	If true, parser will attempt to read events split across line breaks. Default **`false`**	boolean
break_on_endat	.If true, input stream parsing will stop on ENDAT. Default **`false`**	boolean
enable_error_ellipses	If true, will attempt to parser error ellipses from data and use them as event geos. Default **`false`**	boolean
use_unique_identifier_for_track	If true, **`track_number`** event attribute will be parsed from CTC headers. Default **`false`**	boolean

OTHG Event Parser options

2.5.9 WAMI Event Parser

`icg.engine.event.ingest.data.parsers.wami.WAMIEventParser`

Parses WAMI events in XML format.

2.5.10 LUX Json Event Parser

`icg.engine.util.event.parsers.LUXJsonEventParser`

Parses an input JSON String as a LUXEvent. No configuration 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 hostname1:port1,hostname2:port2,hostname3:port3[/chroot/path]	URL

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

2.7 Entity Manager Configuration

Several Analytic plugins, and the EntityManagerEnrichment, have the optional ability to write data to a graph data store. The graph store configuration will be read from `entity_manager.properties`

Property	Description	Data Type or Valid Values
storage_class	The class name for the ERStore to load	One of the classnames of ERStores listed below
<properties for selected graph store>		

Table 94 - ERStoreFactory options

Sample entity_manager.properties

```
storage_class=icg.engine.entity.store.mongo.MongoERStore
mongo_url=localhost
mongo_port=27017
mongo_db=mydb
```

2.7.1 RelationshipConfig

`icg.engine.entity.store.RelationshipConfig`

RelationshipConfig is a convenience class for configuring how Relationships are stored, used in several Analytic plugins.

Property	Description	Data Type or Valid Values
relationship_destination_id_label	This is the key field name in the destination node	String
relationship_destination_type	Type of destination node. This is the	String

	name of the collection containing the destination entity in mongo, typically., “entity”	
relationship_source_id_label	This is the key field name in the source node.	String
relationship_source_type	Type of source node. This is the name of the collection containing the source entity in mongo, typically, “entity”.	String
relationship_type	Type attribute for relationship	String
relationship_description_attribute	Field for store relationship description, in edge collection document	String
relationship_description	Relationship description, identifies origin of relationship.	String
relationship_update_existing	If true, repeated relationships (between the same entities, of the same type), will be updated rather than have new relationships added. Default **`false`**	Boolean

Table 95 - RelationshipConfig options

2.7.2 EntityConfig

`icg.engine.entity.store.EntityConfig`

EntityConfig is a convenience class for configuring how Entities are stored, used in several Analytic and Enrichment plugins.

Property	Description	Data Type or Valid Values
entity_type	Type of Entity	String
entity_key_attribute_label	DB attribute label for key attribute	String

Table 96 - EntityConfig options

2.7.3 MongoERStore

`icg.engine.entity.store.mongo.MongoERStore`

MongoERStore is an entity store manager for the mongo database. It takes the following properties.

Property	Description	Data Type or Valid Values
mongo_url	URL of mongo database server	String (DNS or IP address)
mongo_port	Port on which mongo is listening	Integer, **default 27017**
mongo_username	Authorized mongo username	String
mongo_password	Authorized mongo password	String
mongo_db	Mongo database to access	String
Below properties are optional		

track_collection_suffix	Suffix to entity	String, **default "_tracks"**
track_block_size_millis		Integer, **default 360000**
lod_granularities	Level of detail (lod) granularities	Comma separated integers, decimation factors
lod_input_collection	Should be the entity_tracks collection (use the track_collection_suffix)	String
lod_output_collection	The lod process output (e.g., "entity_tracks_lod")	String
lod_desc_collection	The lod block descriptor collection (e.g., "entity_tracks_lod_desc")	String
lod_entity_type	The type of entry being indexed	String
lod_base_block_size_millis		Integer, **default 1**
lod_block_sizes		Comma separated integers, count must match lod_block_sizes
create_indexes	Whether to automatically create indexes on the collections	Boolean, **default false**
track_expire_after_seconds	Track expiry	Integer, **default never**

MongoERStore options

2.7.4 Entity Manager Enrichment

`icg.engine.enrichment.entity.manager.EntityManagerEnrichment`

Parse Entities and Relationships from data streams, add to events and graph store.

Property	Description	Data Type or Valid Values
entity_evaluator_class.x	Class name for entity evaluator x	**`icg.engine.enrichment.entity.manager.entity.UniqueIDEntityEvaluator`**
entity_evaluator_stream.x	Input stream for entity evaluator x	string
relationship_evaluator_class.x	Class name for relationship evaluator x	**`icg.engine.enrichment.entity.manager.relationship.UniqueIDRelationshipEvaluator`**
relationship_evaluator_stream.x	Input stream for relationship evaluator x	String
location_xpath	Display name for output property	Xpath string

Entity Manager Enrichment options

`icg.engine.enrichment.entity.manager.relationship.UniqueIDRelationshipEvaluator`

Property	Description	Data Type or Valid Values
source_id_xpath	XPath to source ID	XPath String
source_id_storage_label	Label for source ID	String
destination_id_xpath	XPath to dest ID	XPath String

destination_id_storage_label	Label for dest ID	String
relationship_type	Type field for db	String
source_type	Entity type of source	String
destination_type	Entity type of dest	String
attribute_xpath.x	XPath to attribute to store in the relationship	XPath String
attribute_static_value.x	Optional static value for attribute	String
attribute_storage_label.x	DB label for attribute x	String
update	Whether to update an existing relationship or add a new one between the same entities	Boolean

Unique ID Relationship Evaluator options

`icg.engine.enrichment.entity.manager.entity.UniqueIDEntityEvaluator`

Property	Description	Data Type or Valid Values
attribute_xpath.x	XPath to attribute to store in entity	XPath String
attribute_static_value.x	Optional static value for attribute	String
attribute_storage_label.x	Label for attribute	String
attribute_storage_type.x	Attribute type	Long,float,string,other,datestring,geo
key_attribute_xpath.x	XPath to key attribute	XPath String
key_attribute_storage_label.x	Label for key attribute	String
entity_type	Entity type	String
location_xpath	Optional XPath to location	XPath String

Unique ID Entity Evaluator options

Output Property	Description	Data Type
entity_id	ID of detected entity	String
relationship_id	ID of detected relationship	String

Entity Manager Enrichment output properties