

IBM **Guardium**
Discover and Classify

STREAMING API NOTIFICATIONS

IBM GUARDIUM DISCOVER AND CLASSIFY

VERSION 4.2.2

TABLE OF CONTENTS

Table of Contents	2
Introduction	3
What Is Streaming API?	3
Kafka Streaming	3
Subscribing to Kafka Topics	4
IGDC Kafka Topics	6
Confirmed Data Source	6
Database Topology	6
Data Sample	9
System Events	10
Fetching Instructions	11
Incoming Data Sources	12
Confirmed Data Sources	12
Data Asset	13
Data Subject Source	13
Candidate	18
Data Subject Log Records	18
UPR - Unique PII Record	23
Classification Input	24
Data Asset	25
Appliance Info	26
Data Subject Configuration	26
Stats	28
Named Entity Bundle	28
Named Entity Message	30
Tag	31
Discovery Configuration	31
Discovery Command Message	33
Discovery Status Messages	33
Repository Status	35
Data Subject Categories	36
Marketplace Subscription	36

INTRODUCTION

WHAT IS STREAMING API?

Streaming APIs are used to examine data in real-time for users to gather up-to-date information. A streaming API (aka HTTP Push) works by the client opening a socket, providing some criteria of the data it wants to receive and the server will deliver new data as it is received over the open socket.

The streaming API model is usually implemented for reading data. It is used to deliver data to consumers, not to make writes or deletes. The streaming API conveys new data as it is obtained over the open socket, delivering concise results in real-time.

KAFKA STREAMING

The IGDC platform utilizes Apache Kafka - a distributed event streaming platform capable of handling trillions of events a day.



Subscribing to Kafka topics and receiving notifications when a topic has been updated is supported by a variety of platforms and interfaces; For information on how to implement subscription & notification in your system, see [Kafka Documentation](#).

SUBSCRIBING TO KAFKA TOPICS

IGDC uses Kafka to interact between services and provides a way for consumers to subscribe from outside, if firewall rules allow it. This section describes how to connect to Kafka inputs.

In the Kafka Console Consumer tool, execute the command below:

```
kubectl -n {NAMESPACE} exec -t kafka-cm-tls-0 kafka -- kafka-console-consumer.sh --bootstrap-server localhost:9094 --from-beginning --topic sources
```

Table 1: Kafka console consumer tool parameters

VARIABLE	DESCRIPTION
NAMESPACE	Namespace, where IGDC is installed. For example, <i>dc-engine-cm</i> .
kafka-cm-tls-0	Pod name that manages Kafka containers. For example, <i>kafka-cm-tls-0</i> .
kafka-console-consumer.sh	Consumer tool script, which accepts bootstrap-server, topic and other parameters.

Python snippet:

```
from app.settings import *
from confluent_kafka import Consumer

config = {
    'bootstrap.servers': f"{KAFKA_SERVER_HOST}:{KAFKA_SERVER_PORT}",
    'group.id': "sources-1",
    'security.protocol': 'SASL_SSL',
    'sasl.username': KAFKA_USERNAME,
    'sasl.mechanism': 'PLAIN',
    'sasl.password': KAFKA_PASSWORD,
    'enable.auto.commit': True,
    "auto.offset.reset": 'earliest',
    'session.timeout.ms': 30000,
    'ssl.ca.location': f"{KAFKA_CACERT_FILENAME}",
    'ssl.certificate.location': f"{KAFKA_CERT_FILENAME}",
    'ssl.key.location': f"{KAFKA_KEY_FILENAME}",
    'enable.ssl.certificate.verification': False,
    "enable.partition.eof": True
}

consumer = Consumer(config)
consumer.subscribe(['sources'])
while True:
    msg = consumer.consume(1)
    if msg:
        print(msg[0].value())
```

Table 2: Python snippet variables

VARIABLE	DESCRIPTION
KAFKA_SERVER_HOST	Ingress controller public IP address, usually - Node IP: <pre>kubectl get node <node> -o jsonpath='{.status.addresses[?(@.type=="InternalIP")].address}'</pre>
KAFKA_SERVER_PORT	Publicly exposed kafka port for API purposes: <pre>kubectl -n "\$NAMESPACE" get svc kafka-cm-tls-0-external -o jsonpath='{.spec.ports[0].nodePort}'</pre>
KAFKA_	Basic Auth username for Kafka, usually set during deployment:

VARIABLE	DESCRIPTION
USERNAME	<pre>kubectl -n "\$NAMESPACE" get secrets control-plane-kafka -o jsonpath='{.data.username}' base64 -d</pre>
KAFKA_PASSWORD	<p>Basic Auth password for Kafka, also set during deployment:</p> <pre>kubectl -n "\$NAMESPACE" get secrets control-plane-kafka -o jsonpath='{.data.password}' base64 -d</pre>
KAFKA_KEY_FILE	<p>Private key for connection to Kafka:</p> <pre>kubectl -n int-pavlo get secrets control-plane-kafka-tls -o jsonpath='{.data.kafka-0.keystore.key}' base64 -d</pre>
KAFKA_CERT_FILE	<p>PEM file for SSL connection to Kafka:</p> <pre>kubectl -n int-pavlo get secrets control-plane-kafka-tls -o jsonpath='{.data.kafka-0.keystore.pem}' base64 -d</pre>
KAFKA_CACERT_FILE	<p>Truststore for Kafka certificates:</p> <pre>kubectl -n int-pavlo get secrets control-plane-kafka-tls -o jsonpath='{.data.kafka.truststore.pem}' base64 -d</pre>

IGDC KAFKA TOPICS

CONFIRMED DATA SOURCE

TOPIC NAME	confirmed_repositories
DESCRIPTION	Topic is updated when a new data subject or Candidate record/s have been added to or deleted from a data source.
RESPONSE(S)	<pre> { "schema_version": "0.4", "repositoryId": "xxxxxxxxxx2", "applianceId": "xxxxxxxxxxx1", "url": "url", "vendor": "Oracle Couchbase", "type": "Network Cloud Storage Database FS", "workspace": "ltouch", "deleted": true, "has_pii": true, "subjectType": "RDA Candidate" } </pre>

DATABASE TOPOLOGY

TOPIC NAME	dbt_erd
DESCRIPTION	Database topology ERD.
RESPONSE(S)	<pre> { "schema_version": "1.0", "correlationId": "cs1111", "createdAtMs": 90558894322, "id": { "repositoryId": "123e4567-e89b-12d3-a456-426655440000", "schemaId": "s1" }, "deleted": true, "product": "BigQuery", "erd": { "partial": "NO", "tables": [{ "name": "s1.address", "fields": [{ "name": "id", "type": "int4:10" }, { "name": "street", "type": "varchar:2147483647", "entityType": "STREET_ADDRESS" }, { "name": "city", "type": "varchar:2147483647", "entityType": "CITY" }, { "name": "countryId", "type": "int4:10" }] }] } } </pre>

```
    ]},
  {
    "name": "s1.country",
    "fields": [
      {
        "name": "id",
        "type": "int4:10"
      },
      {
        "name": "name",
        "type": "varchar:2147483647",
        "entityType": "COUNTRY"
      }
    ]
  },
  {
    "name": "s1.user_details",
    "fields": [
      {
        "name": "id",
        "type": "int4:10"
      },
      {
        "name": "addressId",
        "type": "int4:10"
      },
      {
        "name": "first_name",
        "type": "varchar:2147483647",
        "entityType": "FIST_NAME"
      },
      {
        "name": "last_name",
        "type": "varchar:2147483647",
        "entityType": "LAST_NAME"
      },
      {
        "name": "sex",
        "type": "varchar:2147483647"
      }
    ]
  },
  {
    "name": "s1.user_details",
    "fields": [
      {
        "name": "id",
        "type": "int4:10"
      },
      {
        "name": "addressId",
        "type": "int4:10"
      },
      {
        "name": "first_name",
        "type": "varchar:2147483647",
        "entityType": "FIST_NAME"
      },
      {
        "name": "last_name",
        "type": "varchar:2147483647",
        "entityType": "LAST_NAME"
      },
      {
        "name": "sex",
```

```
        "type": "varchar:2147483647"
      }
    ]
  }, {
    "name": "s1.users",
    "fields": [
      {
        "name": "id",
        "type": "int4:10"
      },
      {
        "name": "login",
        "type": "varchar:2147483647"
      },
      {
        "name": "email",
        "type": "varchar:2147483647"
      }
    ]
  }
], "refs": [
  {
    "links": [
      {
        "from": "s1.country.id",
        "to": "s1.address.countryId"
      }
    ],
    "type": "AUTO"
  },
  {
    "links": [
      {
        "from": "s1.country.id",
        "to": "1.address.countryId"
      }
    ],
    "type": "AUTO"
  },
  {
    "links": [
      {
        "from": "s1.user_details.id",
        "to": "s1.users.id"
      }
    ],
    "type": "AUTO"
  }
], "idx": [
  {
    "name": "address_pkey",
    "fields": [
      "s1.address.id"
    ],
    "unique": true
  },
  {
    "name": "country_pkey",
    "fields": [
      "s1.country.id"
    ],
    "unique": true
  },
],
```

```

    {
      "name": "user_details_pkey",
      "fields": [
        "s1.user_details.id"
      ],
      "unique": true
    },
    {
      "name": "users_pkey",
      "fields": [
        "s1.users.id"
      ],
      "unique": true
    },
    {
      "name": "idx_address_phone",
      "fields": [
        "s1.users.login",
        "s1.users.email"
      ],
      "unique": false
    }
  ]
}

```

DATA SAMPLE

TOPIC NAME	dbt_data_samples
DESCRIPTION	Displays data samples for each table.
RESPONSE(S)	<pre> { "schema_version": "1.0", "id": { "repositoryId": "123e4567-e89b-12d3-a456-426655440000", "schemaId": "s1" }, "deleted": true, "contentType": "TEXT", "correlationId": "cc1123", "sec": { "type": "token AES", "iv": "1f77bcf86cd279943901", "keyId": "somekey/abc", "keyVer": "1", "applies": ["NATIONAL_ID", "PASSPORT_NUMBER"] } }, "dataSample": [{ "source": "VIEW", "name": "s1.view", "fieldSample": [{ "fieldName": "id", "values": ["1", "2", "3"] }] }] </pre>

```

{
  "source": "TABLE",
  "name": "sl.address",
  "fieldSample": [
    {
      "fieldName": "id",
      "values": ["1", "2", "3"]
    },
    {
      "fieldName": "street",
      "values": ["Burt Ash Rd", "st2", "st2"]
    },
    {
      "fieldName": "city",
      "values": ["London", "Hastings", "Oxford"]
    },
    {
      "fieldName": "countryId",
      "values": ["1", "1", "1"]
    }
  ]
},
{
  "source": "TABLE",
  "name": "sl.country",
  "fieldSample": [
    {
      "fieldName": "id",
      "values": ["1", "2", "3"]
    },
    {
      "fieldName": "name",
      "values": ["UK", "USA", "Ukraine"]
    }
  ]
}
]
}

```

SYSTEM EVENTS

TOPIC NAME	events
DESCRIPTION	Topic for system events.
RESPONSE(S)	<pre> { "schema_version": "0.1", "name": "NewFile", "code": 1000001, "ts": 1000001000, "sourceId": "552f639947d5ce0d79bf841e79fc6b7e", "repositoryId": "552f639947d5ce0d79bf841e79fc6b7e", "applianceId": "552f639947d5ce0d79bf841e79fc6b7e", "storageType": "FS DB NET", "path": "schema.vvid c://xxx/yyy/file.pdf", "contentType": "pdf txt png", "categories": [] } </pre>

FETCHING INSTRUCTIONS

TOPIC NAME	dbt_fetching_instructions
DESCRIPTION	Used to communicate fetching instructions for the database.
RESPONSE(S)	<pre> { "schema_version": "1.0", "id": { "repositoryId": "123e4567-e89b-12d3-a456-426655440000", "schemaId": "s1" }, "product": "BigQuery", "deleted": false, "erdHash": "426655440000e89b12d3", "constraints": [["NATIONAL_ID"], ["PASSPORT_NUMBER"], ["GIVEN_NAME", "PHONE_NUMBER"]], "views": [{ "id": "123e4567-e89b-12d3-a456", "root": "s1.address", "fields": [{ "name": "s1.address.street", "type": "STREET_ADDRESS", "mandatory": true }, { "name": "s1.address.city", "type": "CITY" }, { "name": "s1.country.name", "type": "GIVEN_NAME" }], "refs": [{ "links": [{ "from": "s1.country.id", "to": "s1.address.countryId" }], "type": "AUTO" }], "meta": { "type": "UNCONFIRMED", "changeMethod": "MANUAL", "name": "some optional name", "description": "some optional descr" } }], "instructions": [{ "applyTo": ["s1.country.id", "s1.country.name"], "setProps": { "samples": 100, "trimTo": 256, "models": [{ "id": "calsStats", </pre>

```

        "samples": 1000
      }
    ]
  }
}

```

INCOMING DATA SOURCES

TOPIC NAME	incoming_repositories
DESCRIPTION	Topic for unconfirmed newfound data sources.
RESPONSE(S)	<pre> { "schema_version": "2.0", "applianceId": "some long string", "src": "networkdiscovery discovery manual gdp ...", "location": { "region": { "type": "AWS", "id": "central-west2", "zone": "west-c"}, "address": { "county": "US", "area": "NY/...", "city": "NY", "zip": "2345"}, }, "product": { "product": "ORACLE", "vendor": "ORACLE", //optional. In some cases we should define it from other params "type": "DB" }, "externalIds": ["arn:aws:elasticache:eu-west-3:074851364937:parametergroup: default.memcached1.4"], "params": { "host": "localhost", "port": "27001", "sid": "oracle", "username": "user1", "database": "user2" "owner_id": "some account id" } } </pre>

CONFIRMED DATA SOURCES

TOPIC NAME	confirmed_repositories
DESCRIPTION	Topic for uconfirmed data sources, their description, and ID.
RESPONSE(S)	<pre> { "schema_version": "0.4", "repositoryId": "xxxxxxxxxxx2", "applianceId": "xxxxxxxxxxx1", "url": "url", "vendor": "Oracle Couchbase", "type": "Network Cloud Storage Database FS", "workspace": "1touch", "deleted": true, </pre>

	<pre> "has_pii": true, "subjectType": "RDA Candidate" } </pre>
--	--

DATA ASSET

TOPIC NAME	data_assets
DESCRIPTION	<p>Topic is updated when:</p> <ul style="list-style-type: none"> • A new Data Asset has been created • A Data Asset has been updated • A Data Asset has been deleted.
RESPONSE(S)	<pre> { "type": "UPDATE", "schemaVersion": "2.1", "metadata": { "id": "a8fb62fc-7efa-4dfb-bf8e-7e605e58cd7f", "name": "data asset name", "description": "Customers of Europe, North America, Asia", "sensitivity": "Critical", "type": "PERSONAL SENSITIVE", "dataAssetClassificationType": "CUSTOMER EMPLOYEE ORGANIZATION PARTNER_3RD_PARTY_ORGANIZATION", "businessRepresentative": { "firstName": "John", "lastName": "Connor", "email": "john.connor@skynet.com" }, "businessUnit": "businessUnit", "role": "role", "reasonsOfProcessing": ["reason0", "reason1"], "risks": ["risk1", "risk0"], "linkedRootDataAssets": ["432017797386076449", "432018234453143105", "432017797357764833"], "linkedTags": ["CCPA", "Tokenize", "Move", "LAN"] } } </pre>

DATA SUBJECT SOURCE

TOPIC NAME	sources
DESCRIPTION	<p>Topic is updated when a Source has been added to a Fileshare, Database, or Network Repository. Source can be a File or Record type.</p>

RESPONSE (S)	Example with FS key and File content
	<pre> { "schema_version":"3.0", "id":"552f639947d5ce0d79", "type":"Candidate", "applianceId":"552f6399-47d5-ce0d-79bf-841e79fc6b7e", "lut":1584465351, "key":{ "type":"FS", "repositoryId":"552f6399-47d5-ce0d-79bf-841e79fc6b7e", "path":"/filestorage/important/file.txt" }, "content":{ "type":"File", "fileType":"xls", "meta":{ "size":1875412069075701099, "owner":"John Dow", "lastChanged":"ts", "created":"ts" }, "categories":["341jgBsTA4"], "enttypes":["NATIONID", "..."] } } </pre>
	<pre> { "schema_version":"3.0", "id":"552f6399-47d5-ce0d-79bf-841e79fc6b89", "type":"Candidate", "applianceId":"552f6399-47d5-ce0d-79bf-841e79fc6b7e", "lut":1584465351, "key":{ "type":"Database", "repositoryId":"552f6399-47d5-ce0d-79bf-841e79fc6b7e", "viewId":"schema.vvid" }, "content":{ "type":"Entity", "meta":{ "tableNames":["Table1", "Table2", "Table3"] }, "enttypes":["NATIONID", "..."] } } </pre>
	Example with Database key and File content

```

{
  "schema_version": "3.0",
  "id": "552f6399-47d5-ce0d-79bf-841e79fc6b89",
  "type": "RDA",
  "applianceId": "552f6399-47d5-ce0d-79bf-841e79fc6b7e",
  "lut": 1584465351,
  "key": {
    "type": "Database",
    "repositoryId": "552f6399-47d5-ce0d-79bf-841e79fc6b7e",
    "viewId": "schema.vvid"
  },
  "content": {
    "type": "File",
    "fileType": "xls|doc|txt",
    "meta": {
      "size": 1875412069075701099,
      "owner": "John Dow",
      "lastChanged": "ts",
      "created": "ts"
    },
    "categories": [
      "race|sex|religion"
    ],
    "enttypes": [
      "NATIONID",
      "..."
    ]
  }
}

```

Example with Network key and Entity content

```

{
  "schema_version": "3.0",
  "id": "552f6399-47d5-ce0d-79bf-841e79fc6b89",
  "type": "Candidate",
  "applianceId": "552f6399-47d5-ce0d-79bf-841e79fc6b7e",
  "lut": 1584465351,
  "key": {
    "type": "Network",
    "repositoryId": "552f6399-47d5-ce0d-79bf-841e79fc6b7e",
    "client": "host|ip",
    "path": "fileid|dbid?"
  },
  "content": {
    "type": "Entity",
    "enttypes": [
      "NATIONID",
      "..."
    ]
  }
}

```

	<p>Example with SaaSApplication key and File content</p> <pre> { "schema_version":"3.0", "id":"552f6399-47d5-ce0d-79bf-841e79fc6b89", "type":"Candidate", "applianceId":"552f6399-47d5-ce0d-79bf-841e79fc6b7e", "lut":1584465351, "key":{ "type":"SaaSApplication", "repositoryId":"552f6399-47d5-ce0d-79bf-841e79fc6b7e", "path":"/filestorage/important/file.txt" }, "content":{ "type":"File", "fileType":"txt", "meta":{ "size":1875412069075701099, "owner":"John Dow", "lastChanged":"ts", "created":"ts" }, "categories":["race sex religion"], "enttypes":["NATIONID", "..."] } } </pre>
	<p>Example with SaaSApplication key and Entity content</p> <pre> { "schema_version":"3.0", "id":"552f6399-47d5-ce0d-79bf-841e79fc6b89", "type":"Candidate", "applianceId":"552f6399-47d5-ce0d-79bf-841e79fc6b7e", "lut":1584465351, "key":{ "type":"SaaSApplication", "repositoryId":"552f6399-47d5-ce0d-79bf-841e79fc6b7e", "viewId":"schema.vvid" }, "content":{ "type":"Entity", "meta":{ "tableNames":["Table1", "Table2", "Table3"] }, "enttypes":["NATIONID", "..."] } } </pre>
	<p>Example with Email key and Email content</p>

```

{
  "schema_version": "3.1",
  "id": "552f6399-47d5-ce0d-79bf-841e79fc6b89",
  "type": "Candidate",
  "applianceId": "552f6399-47d5-ce0d-79bf-841e79fc6b7e",
  "lut": 1584465351,
  "key": {
    "type": "Email",
    "repositoryId": "552f6399-47d5-ce0d-79bf-841e79fc6b7e",
    "messageId": "99-47d5-ce0d-79bf-841e79"
  },
  "content": {
    "type": "Email",
    "mailboxOwner": "mailbox.owner",
    "headers": {
      "to": [
        "recp@comp.com",
        "another@email.com"
      ],
      "from": [
        "Name Nick <recp@comp.com>"
      ],
      "sender": [
        "recp@comp.com"
      ],
      "cc": [
        "cc@comp.com"
      ],
      "bcc": [
        "bcc@comp.com"
      ]
    },
    "attachments": [
      "file.xml",
      "file.csv"
    ],
    "folder": "draft|archived|...",
    "messageTs": 1233455,
    "meta": {
      "size": 150000
    },
    "categories": [
      "race|sex|religion"
    ],
    "enttypes": [
      "NATIONID",
      "..."
    ]
  }
}

```

Example with ID key and Entity content

```

{
  "schema_version": "3.3",
  "id": "a4514aeff8786070eb371b6370f4ce3e204f5de12713ee783cf83819c8f9a622",
  "type": "Candidate",
  "applianceId": "1e278399-56d5-ab45-79bf-841e79fc54e2",
  "lut": 13491823,
  "key": {
    "type": "ID",
    "repositoryId": "552f6399-47d5-ce0d-79bf-841e79fc6b7e",

```

```

        "id": "D011S26RA10",
        "details": {
            "name": "upl",
            "link": "link"
        }
    },
    "content": {
        "type": "Entity",
        "enttypes": [
            "TAX_ID",
            "CC_NUMBER"
        ]
    }
}
    
```

CANDIDATE

TOPIC NAME	pii_candidates
DESCRIPTION	Topic is updated when a new Candidate record has been identified. The response will include all data element data associated with the Candidate record.
RESPONSE(S)	<pre> { "schema_version": "1.1", "id": "1335869f90ealc664f6ecc1deab97f9e37c8ad27d8e136b931ab16e16022bb56", "piiSrcId": "507f1f77bcf86cd799439012", "meta": { "cycleId": "6cd97f1012507997bc43f7f8" }, "personalInfo": { "NATIONAL_ID": ["279-87-6784"], "PASSPORT_NUMBER": ["629627248"], "DRIVER_LICENSE": ["G776208445599"], "TAX_ID": ["392655084009"], "CC_NUMBER": ["2434392655084637"], "CC_TYPE": ["Mastercard"], "CVV2": ["311"], "CC_EXPIRES": ["4/2019"], "GIVEN_NAME": ["Jennifer"], "VEHICLE_NUMBER": ["ZHV-421"], "PHONE_NUMBER": ["771-658-2269"], "SURNAME": ["Pelegrin"], "BIRTHDAY": ["03/12/99"], "CITY": ["Anchorage"], "STREET_ADDRESS": ["2189 Timbercrest Road"], "GENDER": ["Female"], "FULL_NAME": ["Ms. Jennifer Pelegrin"], "MOTHERS_MAIDEN": ["Khaleel"], "COUNTRY_CODE": ["USA"], "COUNTRY": ["United States"], "STATE": ["Alaska"], "ZIP_CODE": ["20428"], "EMAIL_ADDRESS": ["pelegrin@cutekittens.com"], "USERNAME": ["13khaleel9"] } } </pre>

DATA SUBJECT LOG RECORDS

TOPIC NAME	pii_log_record
------------	-----------------------

DESCRIPTION	Topic where crawlers send candidates in the form of log records.
RESPONSE (S)	FS-File example <pre>{ "schema_version": "3.1", "type": "Candidate", "cycleId": "79bf-841e79fc6b7e552f639947d5ce0d", "source": { "applianceId": "552f6399-47d5-ce0d-79bf-841e79fc6b7e", "ts": 1123344455, "key": { "type": "FS", "repositoryId": "552f6399-47d5-ce0d-79bf-841e79fc6b7e", "path": "/important/personalData.pdf", "originDetails": { "url": "url", "vendor": "xxxxxxx", "repositoryType": "CloudStorage" } } }, "content": { "type": "File", "fileType": "pdf", "meta": { "size": 150000, "owner": "John Dow", "lastChanged": 12334, "created": 1234 }, "categories": [] }, "personalInfo": [{ "NATIONAL_ID": ["709-48-1890"], "CC_NUMBER": ["xxxxxxx"] }] }</pre>

DB-Ent example

```
{
  {
    "schema_version":"3.1",
    "type":"Candidate",
    "cycleId": "79bf-841e79fc6b7e552f639947d5ce0d",
    "source": {
      "applianceId": "552f6399-47d5-ce0d-79bf-841e79fc6b7e",
      "ts": 1123344455,
      "key":{
        "type":"Database",
        "repositoryId":"552f6399-47d5-ce0d-79bf-841e79fc6b7e",
        "viewId":"552f6399-47d5-ce0d-79bf-841e79fc6b7e",
        "originDetails": {
          "url": "url",
          "vendor": "vendor",
          "repositoryType": "Database"
        }
      }
    },
    "content":{
      "type":"Entity",
      "meta": {"tableNames": ["users","emails",""]}
    }
  },
  "personalInfo":[
    {
      "NATIONAL_ID":["709-48-1890"],
      "CC_NUMBER":["4765xxxxxxxx2871"]
    }
  ]
}
```

DB-File example

```

{
  "schema_version":"3.1",
  "type":"Candidate",
  "cycleId": "79bf-841e79fc6b7e552f639947d5ce0d",
  "source": {
    "applianceId": "552f6399-47d5-ce0d-79bf-841e79fc6b7e",
    "ts": 1123344455,
    "key":{
      "type":"Database",
      "viewId":"552f6399-47d5-ce0d-79bf-841e79fc6b7e",
      "repositoryId": "xe",
      "originDetails": {
        "url": "url",
        "vendor": "xxxxxx",
        "repositoryType": "Database"
      }
    }
  },
  "content":{
    "type":"File",
    "fileType": "pdf",
    "meta": {"size": 1234445},
    "categories": ["political"]
  }
},
"personalInfo":[
  {
    "NATIONAL_ID":["709-48-1890"],
    "CC_NUMBER":["4765xxxxxxxx2871"]
  }
]
}

```

Network-Entity

```

{
  "schema_version":"3.1",
  "type":"Candidate",
  "cycleId": "79bf-841e79fc6b7e552f639947d5ce0d",
  "source": {
    "applianceId": "552f6399-47d5-ce0d-79bf-841e79fc6b7e",
    "ts": 1223334,
    "key":{
      "type":"Network",
      "originDetails":{
        "url":"http://somehost:8890",
        "vendor":"centeraCMS",
        "client":"clienthostname",
        "repositoryType": "Central Storage"
      }
    }
  },
  "content":{
    "type": "Entity"
  }
},
"personalInfo":[
  {
    "NATIONAL_ID":["709-48-1890"],
    "CC_NUMBER":["xxxxxxx"]
  }
]
}

```

Salesforce - entity

```
{
  "schema_version": "3.1",
  "type": "RDA",
  "cycleId": "79bf-841e79fc6b7e552f639947d5ce0d",
  "source": {
    "applianceId": "552f6399-47d5-ce0d-79bf-841e79fc6b7e",
    "ts": 1617791241240,
    "key": {
      "type": "Database",
      "viewId": "fc6b7e552f639947d5",
      "repositoryId": "RDA:552f6399-47d5-ce0d-79bf-841e79fc6b7e",
      "originDetails": {
        "url": "https://dsafsdfsdf.salesforce.com",
        "vendor": "salesforce",
        "repositoryType": "SaaS Application"
      }
    }
  },
  "content": {
    "type": "Entity",
    "meta": {
      "tableNames": ["User"]
    }
  }
},
"personalInfo": [
  {
    "NATIONAL_ID": ["709-48-1890"],
    "CC_NUMBER": ["xxxxxxx"]
  }
]
}
```

Salesforce - unstructured data	<pre> { "schema_version": "3.1", "type": "Candidate", "cycleId": "79bf-841e79fc6b7e552f639947d5ce0d", "source": { "applianceId": "552f6399-47d5-ce0d-79bf-841e79fc6b7e", "ts": 1123344455, "key": { "type": "FS", "repositoryId": "552f6399-47d5-ce0d-79bf-841e79fc6b7e", "path": "ContentVersion:552f639947d5", //or https://12345-app.- //salesforce.com/services/data/v50.0/subjects/ContentVersion/552f639947d5 "originDetails": { "url": "https://12345-app.salesforce.com", "vendor": "salesforce", "repositoryType": "SaaS Application" } } }, "content": { "type": "File", "fileType": "pdf", "meta": { "size": 150000, "owner": "John Dow", "lastChanged": 12334, "created": 1234 }, "categories": [] } }, "personalInfo": [{ "NATIONAL_ID": ["709-48-1890"], "CC_NUMBER": ["xxxxxxx"] }] } </pre>
--------------------------------	---

UPR - UNIQUE PII RECORD

TOPIC NAME	upr
DESCRIPTION	Topic is updated when a new confirmed data subject copy has been identified. The response will include all data element data associated with the data subject record.
RESPONSE(S)	<pre> { "schema_version": "0.1", "id": "507f1f77bcf86cd799439011", "deleted": true, "rda_src": ["507f1f77bcf86cd799439011", "507f1f77bcf86cd799439011"], "personalInfo": { "NATIONAL_ID": [{ "value": "ab777-555-xx", "refs": ["507f1f77bcf86cd799439012", "507f1f77bcf86cd799439013", "507f1f77bcf86cd799439014", "507f1f77bcf86cd799439018"], "lut": 1584465351 }] } } </pre>

```

        {
            "value": "ab777-555-yy",
            "refs": ["507f1f77bcf86cd799439012"],
            "lut": 1584465351
        }
    ],
    "PASSPORT_NUMBER": [
        {
            "value": "XXX-YYY-7777",
            "refs": ["507f1f77bcf86cd799439014"],
            "lut": 1584465351
        }
    ],
    "DRIVER_LICENSE": [],
    "TAX_ID": [],
    "CC_NUMBER": []
}
    }
}

```

CLASSIFICATION INPUT

TOPIC NAME	classification-input
DESCRIPTION	Topic for document content.
RESPONSE(S)	File type
	<pre> { "txt": "sample text", "source": { "sourceType": "FILE", "repository": { "type": "CENTRAL_STORAGE", "vendor": "CIFS_GENERAL", "path": "Yuri", "networkElement": { "hostname": "172.16.3.141", "ip": null, "type": "NETWORK_DEVICE", "topology": null } } }, "name": "Pirkko2.pptx", "path": "/", "created": 1580394384768, "type": "files", "lastChanged": 1580394385456, "size": 381671 } </pre>

Transaction	<pre> { "txt": "El 25 de julio, en la cima de la avenida Mazepa, un automóvil Chevrolet Bolt, conducido por un conductor ebrio Thomas Smith, entró en un minibús de pasajeros de la ruta 156. Después de un tiempo, el conductor del automóvil desapareció de la escena de un accidente. La solicitud de todos los que vieron o conocen la ubicación exacta del automóvil buscado para informar al jefe de los sindicatos de automovilistas Mikhail Tonkonogom o la policía. La matrícula del culpable es 933-ORZ, frente al automóvil está esencialmente rota y no podía pasar desapercibido muy lejos.", "source" : { "sourceType" : "TRANSACTION", "protocol" : "HTTP", "source" : { "hostname" : "JU8PU3RS1", "ip" : "5aBGW731", "type" : "NETWORK_DEVICE", "topology" : "LAN" }, "destination" : { "hostname" : "A8T81", "ip" : "3", "type" : "SQL_SERVER", "topology" : "EXTERNAL" }, "document" : { "name" : "80ow8UKM5MHxB", "created" : 1629430351517, "type" : "json", "checkSum" : 1810927428, "categories" : ["Lj"] }, "ts" : 1453459805361 } </pre>
-------------	--

DATA ASSET

TOPIC NAME	data_assets
DESCRIPTION	Dictionary topic for data assets. The topic is updated when a new Data Asset has been created or a Data Asset has been updated or deleted.
RESPONSE(S)	<pre> { "type": "UPDATE", "schemaVersion": "2.1", "metadata": { "id": "a8fb62fc-7efa-4dfb-bf8e-7e605e58cd7f", "name": "data asset name", "description": "Customers of Europe, North America, Asia", "sensitivity": "Critical", "type": "PERSONAL SENSITIVE", "dataAssetClassificationType": "CUSTOMER EMPLOYEE ORGANIZATION PARTNER_3RD_PARTY_ORGANIZATION", "businessRepresentative": { "firstName": "John", "lastName": "Connor", "email": "john.connor@skynet.com" }, "businessUnit": "businessUnit", "role": "role", </pre>

```

    "reasonsOfProcessing": [
      "reason0",
      "reason1"
    ],
    "risks": [
      "risk1",
      "risk0"
    ],
    "linkedRootDataAssets": [
      "432017797386076449",
      "432018234453143105",
      "432017797357764833"
    ],
    "linkedTags": [
      "CCPA",
      "Tokenize",
      "Move",
      "LAN"
    ]
  }
}

```

APPLIANCE INFO

TOPIC NAME	appliance
DESCRIPTION	Topic for appliance ID and Name.
RESPONSE(S)	<pre> { "schema_version": "1.0", "applianceId": "bb678077-7506-5919-9a8b-74ec9a45745d", "applianceDesc": "Test appliance", "location": { "region": { "type": "AWS", "id": "central-west2", "zone": "west-c" }, "address": { "country": "US", "area": "NY/...", "city": "NY", "zip": "2345" } } } </pre>

DATA SUBJECT CONFIGURATION

TOPIC NAME	pii_configuration
DESCRIPTION	Topic for configuration for Data Subject entities.
RESPONSE(S)	<pre> { "schema_version": "1.0", "available": [{ "locale": { "id": "en-US", "language": "English", "country": "United States" }, "entities": ["BIRTHDAY", "VEHICLE_NUMBER", "PASSPORT_NUMBER", "DRIVER_LICENSE"] }] }, </pre>

```
{
  "locale": {
    "id": "en-GB",
    "language": "English",
    "country": "United Kingdom"
  },
  "entities": [
    "BIRTHDAY",
    "VEHICLE_NUMBER",
    "PASSPORT_NUMBER",
    "DRIVER_LICENSE"
  ]
},
"configured": {
  "locales": [
    "en-US",
    "en-GB"
  ],
  "entities": [
    {
      "name": "BIRTHDAY",
      "indexed": true
    }, {
      "name": "VEHICLE_NUMBER",
      "indexed": true
    }, {
      "name": "VEHICLE_TYPE",
      "indexed": false
    }, {
      "name": "EMAIL_ADDRESS",
      "indexed": true
    }, {
      "name": "NATIONAL_ID",
      "indexed": true
    }, {
      "name": "PASSPORT_NUMBER",
      "indexed": true
    }
  ],
  "constraints": [
    [
      "NATIONAL_ID"
    ],
    [
      "PASSPORT_NUMBER"
    ],
    [
      "DRIVER_LICENSE"
    ],
    [
      "TAX_ID"
    ],
    [
      "CC_NUMBER"
    ],
    [
      "GIVEN_NAME",
      "VEHICLE_NUMBER"
    ],
    [
      "GIVEN_NAME",
      "PHONE_NUMBER"
    ]
  ]
}
```

```

        ],
        [
            "GIVEN_NAME",
            "SURNAME",
            "BIRTHDAY"
        ],
        [
            "GIVEN_NAME",
            "SURNAME",
            "CITY",
            "STREET_ADDRESS"
        ]
    ]
}
}

```

STATS

TOPIC NAME	stats
DESCRIPTION	Common topic where component may publish stats useful for other components.
RESPONSE(S)	<pre> { "schema_version": "0.1", "messageId": "your-app-name_timestamp", "type": "yourStatisticsType", "data": [// any data structure] } </pre>

NAMED ENTITY BUNDLE

TOPIC NAME	named_entity_bundle
DESCRIPTION	List of all data elements locales and matchers.
RESPONSE(S)	<pre> { "schema_version": "2.0", "locales": [{ "id": "en-US", "language": "English", "country": "United States" }, { "id": "en-GB", "language": "English", "country": "United Kingdom" },], "matchers": [{ "id": "onetouch.functions.isAlphabetic" }, { "id": "onetouch.functions.isAlphanumeric" }, { "id": "onetouch.functions.isBrazilBankAccount" }, { "id": "onetouch.functions.maxSize", "args": ["uint"], "title": "Checks the maximum size of the expression" }] </pre>

```

    }, {
      "id": "onetouch.dictionary.cc_type.en"
    }, {
      "id": "onetouch.keywords.cc_type.en"
    }, {
      "id": "onetouch.keywords.cc_type.es-mx"
    }, {
      "id": "onetouch.keywords.cc_type.iw-il"
    }, {
      "id": "onetouch.patterns.phone_number.pt-br"
    }, {
      "id": "onetouch.patterns.postcode.en-gb"
    }, {
      "id": "onetouch.patterns.postcode.iw-il"
    }, .....
  ],
  "entity-packs": [{
    "id": "en-GB",
    "entities": [{
      "id": "ACCOUNT_NUMBER",
      "rules": {
        "and": [{
          "or": ["onetouch.keywords.account_number.en"],
          {
            "or": ["onetouch.patterns.account_number.en"]
          }, {
            "and": ["onetouch.functions.isNumeric", {
              "onetouch.functions.minSize": ["6"],
              {
                "onetouch.functions.maxSize": ["26"]
              }
            ]
          }
        ]
      }
    }, {
      "id": "BIRTHDAY",
      "rules": {
        "and": [{
          "or": ["onetouch.keywords.birthday.en"],
          {
            "or": ["onetouch.patterns.date"]
          }, {
            "and": ["onetouch.functions.isAlphanumeric", {
              "onetouch.functions.minSize": ["6"],
              {
                "onetouch.functions.maxSize": ["17"]
              }
            ]
          }
        ]
      }
    }, .....
  ]
}
]
}

```

NAMED ENTITY MESSAGE

TOPIC NAME	named_entity_definitions
DESCRIPTION	Data element definition.
RESPONSE(S)	<pre> { "schema_version": "2.0", "id": "NATIONAL_ID", "name": "National ID", "version": "bd90ac36-b272-478a-a5a0-3d6beb22fadf", "in_use": true, "description": "A national identification number, national identity number, or national insurance number is used by the governments of many countries as a means of tracking their citizens, permanent residents, and temporary residents for the purposes of work, taxation, government benefits, health care, and other governmentally-related functions", "rules": { "and": [{ "or": ["onetouch.keywords.national_id.en-gb", "custom.keywords.1"] }, { "or": ["onetouch.patterns.national_id.en-gb", "custom.patterns.1"] }, { "or": ["custom.dictionary.1"] }, { "and": ["custom.functions.isNationalId" , { "onetouch.functions.minSize": ["9"] }, { "onetouch.functions.maxSize": ["13"] }] }] }, "definitions": { "keywords": [{ "id": "custom.keywords.1", "content": ["keyword1", "keyword2", "keywordn .. n"], "confidence": 80 }], "dictionaries": [{ "id": "custom.dictionary.1", "content": ["word1", "word1", "word .. n"] }], "patterns": [{ "id": "custom.patterns.1", "content": ["**(a-z){1,3}", "[A-Za-z0-9]{8}-[A-Za-z0-9]{4}-[A-Za-z0-9]{4}-[A-Za-z0-9]{4}-[A-Za-z0-9]{12}"] }] } </pre>

TAG

TOPIC NAME	tags
DESCRIPTION	Tags distribution topic.
RESPONSE(S)	<p>Create tag sample message.</p> <pre> { "schema_version": "0.3", "tag": { "id" : "tag1", "meta": { "type": "type", "description": "description" } } } </pre>
	<p>Delete tag sample message.</p> <pre> { "schema_version": "0.3", "tag": { "id" : "tag1", "deleted" : true } } </pre>

DISCOVERY CONFIGURATION

TOPIC NAME	discovery_configure
DESCRIPTION	Configuration of the discovery remote module.
RESPONSE(S)	<pre> { "schema_version": "1.0", "type": "Configure", "applianceId": "bb678077-7506-5919-9a8b-74ec9a45745d", "configurationId": "bb678077-1111-2222-3333-234253432222", "ts": 12312321324, "DiscoverySettings": { "domainListSettings": [{ "id": 0, "domain": "1touch.corp" }, { "id": 1, "domain": "1touch.io" }], "ipRangeListSettings": [{ "id": 0, "type": "SUBNET", "subnet": "1.1.1.1/16", "ipFrom": "1.1.0.1", "ipTo": "1.1.255.254", "networkLocation": "DMZ" }], } } </pre>

```
{
  "id": 1,
  "type": "IP_RANGE",
  "subnet": "",
  "ipFrom": "10.10.10.1",
  "ipTo": "10.10.10.10",
  "networkLocation": "LAN"
},
"dnsSettings": {
  "primaryDns": "8.8.8.8",
  "secondaryDns": "8.8.8.8",
  "tertiaryDns": "8.8.8.8",
  "domainSearch": null,
  "maxQueries": 10,
  "skipDns": false,
  "queryTimeout": 10
},
"networkTrafficSettings": {
  "typeMechanism": "HARDWARE_ACCELERATED",
  "coreIdMin": 1,
  "coreIdMax": 3,
  "memoryPageSize": 2048,
  "pagesCount": 4096,
  "memoryChannels": 2,
  "snapshotLength": 0,
  "mtuSize": 9600,
  "capturePool": 400000,
  "captureLimit": 200000,
  "captureTimeout": 20000,
  "maxCaptureFilterSize": 1000,
  "ignoreAddresses": [
    ""
  ],
  "additionalDrivers": [
    "librte_pmd_vmxnet3_uio.so",
    "librte_mempool_ring.so"
  ],
  "emulatorPcapPath": "",
  "pcapPortsNumber": 1,
  "pcapsList": [
    ""
  ]
},
"interfacesSettings": {
  "interfaces": [
    "41:00:0"
  ]
},
"networkTrafficLogSetting": {
  "driverLogLevel": 1,
  "logCategoryInitialization": true,
  "logCategoryInitPortsInfo": true,
  "logCategoryJNIWraps": true,
  "logCategoryEvents": false,
  "logCategoryBlacklistPackets": false,
  "logCategoryBlacklistConfig": false,
  "logCategoryCaptPackets": true,
  "logCategoryCaptPacketsEveryPacket": false,
  "logCategoryCaptPacketsFiltered": false,
  "logCategoryCaptPacketsStatistics": true,
  "logCategoryCaptPacketsDiskOps": false,
  "logCategoryEal": true,
}
```

```

    "logCategoryCaptPacketsSkipped": false,
    "logCategoryDebugProtocols": false
  },
  "networkProtocolsSetting": {
    "smb": true,
    "tns": true,
    "tds": true,
    "http": true,
    "mysql": true,
    "couchbase": true,
    "pgsql": true,
    "cql": true,
    "tls": true,
    "kafka": true,
    "mongo": true,
    "smtp": true,
    "ftp": true,
    "db2": true,
    "nfs": true,
    "udp": true,
    "ipLogFiltering": "",
    "portLogFiltering": "",
    "protocolLogFiltering": ""
  },
  "startOnBoot": true,
  "productionMode": true,
  "protocolRegistryFile": "protocols-registry.json"
}

```

DISCOVERY COMMAND MESSAGE

TOPIC NAME	discovery_actions
DESCRIPTION	Performing actions on the discovery remote module.
RESPONSE(S)	<pre> { "schema_version": "1.0", "type": "Actions", "applianceId": "bb678077-7506-5919-9a8b-74ec9a45745d", "commandId": "bb678077-1111-2222-3333-234253432222" "ts": "command": "Start" / "Stop" / "Bind" / "Unbind" "commandArgs": {} } </pre>

DISCOVERY STATUS MESSAGES

TOPIC NAME	discovery_status
DESCRIPTION	Sending from the discovery remote module status and statistics for CM needs.
RESPONSE(S)	<pre> { "schemaVersion": "1.0", "type": "Status", "applianceId": "bb678077-7506-5919-9a8b-74ec9a45745d", "configurationId": "bb678077-1111-2222-3333-234253432222", "ts": 12312321324, "operationMode": "Running", "StatisticsAndStatus": { </pre>

```
"uptime": 1011111,  
"cycleNumber": 245,  
"interfaces": [  
  {  
    "interfaceId": 0,  
    "pci": "80:00.0",  
    "status": "up",  
    "state": "binded",  
    "trafficStats": {  
      "rxPacketsTotal": 1231,  
      "rxPacketsProcessed": 1000,  
      "rxPacketsFiltered": 0,  
      "rxBytesTotal": 12314654,  
      "rxBytesFiltered": 123456,  
      "rxMisses": 3,  
      "rxErrors": 0,  
      "rxNoBufErrors": 0  
    },  
    "discoveredStats": {  
      "foundPii": 100,  
      "foundRepositories": 10,  
      "foundNetworkElement": 1000,  
      "foundFiles": 50,  
      "sentDocumentForProcessing": 20  
    }  
  },  
  {  
    "interface": 1,  
    "pci": "80:00.1",  
    "status": "down",  
    "state": "binded",  
    "trafficStats": {  
      "rxPacketsTotal": 0,  
      "rxPacketsProcessed": 0,  
      "rxPacketsFiltered": 0,  
      "rxBytesTotal": 0,  
      "rxBytesFiltered": 0,  
      "rxMisses": 0,  
      "rxErrors": 0,  
      "rxNoBufErrors": 0  
    },  
    "discoveredStats": {  
      "foundPii": 100,  
      "foundRepositories": 10,  
      "foundNetworkElement": 1000,  
      "foundFiles": 50,  
      "sentDocumentForProcessing": 20  
    }  
  },  
  {  
    "interface": 2,  
    "pci": "80:00.2",  
    "status": "down",  
    "state": "unbinded",  
    "trafficStats": {  
      "rxPacketsTotal": 0,  
      "rxPacketsProcessed": 0,  
      "rxPacketsFiltered": 0,  
      "rxBytesTotal": 0,  
      "rxBytesFiltered": 0,  
      "rxMisses": 0,  
      "rxErrors": 0,  
      "rxNoBufErrors": 0  
    }  
  }  
]
```

```

    },
    "discoveredStats": {
      "foundPii": 100,
      "foundRepositories": 10,
      "foundNetworkElement": 1000,
      "foundFiles": 50,
      "sentDocumentForProcessing": 20
    }
  }
]
}
}

```

REPOSITORY STATUS

TOPIC NAME	repository_status
DESCRIPTION	Current information about repository from the RM point of view.
RESPONSE(S)	<pre> { "schema_version": "1.0", "id": "xxx111222", "applianceId": "", "url": "url", "type": "Database Cloud Storage Cloud Application Email", "vendor": "Oracle", "creationDate": 12234453, "modified": 3345565, "origin": "manual automatic", "deleted": false, "lastJob": { "id": "", "status": "NEW FAILED RUNNING...", "started": 122334, "created": 334455, "duration": 333333 }, "tags": [{ "name": "tag1", "type": "USER_TAG" }, { "name": "tag2", "type": "USER_TAG" }, { "name": "tag1", "type": "DATA_ASSET_TAG" }], "options":[""], "currentStrategy": "", "stats": { "totalScans": 4, "successfulScans": 4 } } </pre>

DATA SUBJECT CATEGORIES

TOPIC NAME	pii_categories
DESCRIPTION	Storage (CRUD) for data subject categories.
RESPONSE(S)	<pre>{ "schema_version": "1.0", "deleted": true, "name": "category name as ID", "description": "not required category description" }</pre>

MARKETPLACE SUBSCRIPTION

TOPIC NAME	marketplace_subscription
DESCRIPTION	Data subject sources and target integration description
RESPONSE(S)	<pre>{ "message": { "origin": "sources", "body": { "schema_version": "3.0", "id": "552f639947d5ce0d79", "type": "Candidate", "applianceId": "552f6399-47d5-ce0d-79bf-841e79fc6b7e", "lut": "1584465351", "key": { "type": "FS", "repositoryId": "552f6399-47d5-ce0d-79bf-841e79fc6b7e", "path": "/filestorage/important/file.txt" }, "content": { "type": "File", "fileType": "xls", "meta": { "size": "1875412069075701099", "owner": "John Dow", "lastChanged": "ts", "created": "ts" } }, "categories": ["341jgBsTA4"], "enttypes": ["NATIONID", "..."] } }, "target": [{ "solution": "mip", "action": "alert", "configurations": { "message_syntax": "" } }] }</pre>

IBM, the IBM logo, and IBM Guardium Discover and Classify are trademarks or registered trademarks of International Business Machines Corporation, in the United States and/or other countries. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on ibm.com/trademark.