The Default TAXII Query Specification

Version 1.0 RC1

Mark Davidson, Charles Schmidt

12/20/2013

The Trusted Automated eXchange of Indicator Information (TAXII™) specifies mechanisms for exchanging structured cyber threat information between parties over the network. This document describes how to express TAXII messages using an XML binding.

Trademark Information

TAXII is a trademark of The MITRE Corporation.

This technical data was produced for the U. S. Government under Contract No. HSHQDC-11-J-00221, and is subject to the Rights in Technical Data-Noncommercial Items clause at DFARS 252.227-7013 (NOV 1995)

©2013 The MITRE Corporation. All Rights Reserved.

Feedback

Feedback on this or any of the other TAXII specifications is welcome and can be sent to taxii-discussion-list@lists.mitre.org after signing up on the community registration page (http://taxii.mitre.org/community/registration.html). You may also provide feedback directly to MITRE by sending a message to taxii@mitre.org.

Comments, questions, suggestions, and concerns are all appreciated.

Table of Contents

1   Introduction

The TAXII Services Specification 1.1 defines the TAXII Query capability, which is an extension point within TAXII. This document defines the Default TAXII Query, which is one implementation of the TAXII 1.1 Query extension point.

1.1   The Default TAXII Query Specification

This specification defines the Default TAXII Query, which is one extension of TAXII Query. As required by the TAXII Services Specification, this document defines structures to be used for TAXII Query (the Query Structure and Query Information Structure) as well as semantics and workflows for processing those structures.

The Default TAXII Capability Specification defines the Default TAXII Query structure, processing rules for the Default TAXII Query, an XML representation of the Default TAXII Query structure to be used in conjunction with the TAXII 1.1 XML Message Binding, and concepts fundamental to the Default TAXII Query.

1.1.1   TAXII Query Format ID for XML

The TAXII Query Format ID for the version of the Default TAXII Query described in this specification is:

$$\texttt{urn:taxii.mitre.org:query:default:1.0}$$

1.2   Document Conventions

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this specification are to be interpreted as described in IETF RFC 2119. [3]

1.3   Terms and Definitions

This document uses the Terms and Definitions defined in the TAXII Services Specification and TAXII Overview [4]. In addition, this document defines terms that are assigned a specific meaning within this specification.

1.3.1   Default TAXII Query Terms

Capability Module – A defined set of relationships (e.g., equals, greater than) that can be used in specifying selection criteria.

Targeting Expression – An expression that specifies the target region of a record for searching.

Targeting Expression Vocabulary – A defined set of vocabulary items to be used in a Targeting Expression.

Node – One vocabulary item in a Targeting Expression Vocabulary.

2 TAXII Default Query

TAXII Default Query allows a Consumer to provide a Producer with selection criteria to use when fulfilling requests for data from a TAXII Data Collection. This section defines The TAXII Default Query.

2.1 Query Structure

The following table details the query structure of the Default Query Structure. This structure is used within the Query field of a Poll Request and the Query field of a Manage Collection Subscription Request with an Action of SUBSCRIBE. This structure contains the criteria that content should be evaluated against when fulfilling a subscription or Poll Request.

Table 1 – Default Query Structure

			Name	Required?	Multiple?	Description
Default Query						This field contains the query.
	Targeting Expression Vocabulary ID			Yes	No	Identifies the Target Expression Vocabulary used in this query. All Targets MUST use only the identified vocabulary.
	Criteria			Yes	No	This field contains the criteria
		Operator		Yes	No	AND / OR - AND indicates that this Criteria evaluates to True if and only if all child Criteria and Criterion evaluate to True. - OR indicates that this Criteria evaluates to True if any child Criteria or Criterion evaluate to True.
		Criteria		At least one of either. Can be multiple of both. All criteria must appear before all criterion.	Yes	The element name indicates the message body type. Its body MUST consist only of the indicated XML Fields.
		Criterion			Yes	This field contains the criterion.
			Negate	No	No	Indicates whether the final result of the Criterion should be negated. If absent, treat this field as "false".
			Target	Yes	No	Contains the Targeting Expression for this Criterion, identifying the region of the record that is being targeted. The Targeting Expression MUST only use Nodes from the specified Target Expression Vocabulary ID.
			Test	Yes	No	Contains the test for the region of the record identified by the Target.

Name					Required?	Multiple?	Description
			Capability ID		Yes	No	Contains the Capability ID, which identifies a Capability Module.
			Relationship		Yes	Yes	Contains the relationship. This value MUST be defined by the Capability Module identified by the Capability ID.
			Parameter		-	-	Contains the parameter(s) for this test. Whether a parameter is required, and whether multiple are permitted is defined by the Capability Module.
				Name	Yes	No	Contains the name of the parameter.

2.1.1 XML Representation

This section defines the XML representation of the Query Structure. This structure is intended for use with the TAXII XML Message Binding 1.1 (urn:taxii.mitre.org:message:xml:1.1).

The XML Namespace for this representation is: http://taxii.mitre.org/query/taxii_default_query-1

Table 2 - XML Representation of TAXII Default Query

XML Name				Data Model Name	#	Description
`<DefaultQuery>`				Default Query	1	The element name indicates that this is a query. Its body MUST consist of only the indicated fields.
	`@targeting_expression_id`			Targeting Expression ID	1	An XML AnyURI.
	`<Criteria>`			Criteria	1	An XML element. This element MUST consist only of the indicated XML fields.
		`@operator`		Operator	1	An XML string restricted to two choices: • OR • AND
		`<Criteria>`		Criteria	1-n	An XML element. This element MUST consist only of the indicated XML fields.
		`<Criterion>`		Criterion		An XML element. This element MUST consist only of the indicated XML fields.
			`@negate`	Negate	0-1	An XML boolean. The default value for this field is 'false'.
			`<Target>`	Target	1	An XML string.
			`<Test>`	Test	1	An XML element. This element MUST consist only of the indicated XML fields.

6

XML Name				Data Model Name	#	Description
			`@cap abil ity_ id`	Capability ID	1	An XML AnyURI.
			`@rel atio nshi p`	Relationshi p	1	An XML string.
			`<Par amet er>`	Parameter	0-n	An XML string.
				`@n am e`	Name	1	An XML string.

2.2 Query Information Structure

The following table details the query structure of the Default Query Information Structure. This structure is used within the Supported Query field of a Discovery Response.

Table 3 - Default Query Information Structure

Name		Required?	Multiple?	Description
Default Query Information		Yes	No	Contains the query information. This field indicates which Targeting Expressions and Capability Modules are supported.
	Targeting Expression ID	Yes	Yes	Indicates a supported Targeting Expression.
	Capability Module	Yes	Yes	Contains a Capability Module ID, indicating a supported Capability Module. This may be a Capability Module defined by this specification or by a third party.

2.2.1 XML Representation

This section defines the XML representation of the Query Structure. This structure is intended for use with the TAXII XML Message Binding 1.1 (urn:taxii.mitre.org:message:xml:1.1).

The XML Namespace for this representation is: http://taxii.mitre.org/query/taxii_default_query-1

XML Name	Data Model Name	Multiple ?	Description
`<DefaultQue ryInformati on>`	Default Query Information	1	The element name indicates that this is a query information structure. Its body MUST consist of only the indicated fields.

7

XML Name	Data Model Name	Multiple ?	Description
`<Targeting Expression Id>`	Targeting Expression ID	1-n	An XML AnyURI.
`<Capabilit yModule>`	Capability Module	1-n	An XML AnyURI.

2.3 Query Evaluation

This section defines how queries are evaluated.

When a Query structure is present, the consumer is requesting only the records from a TAXII Data Collection that meet the specified criteria. If a Query is present and the producer is incapable or unwilling to process the Query, the producer should indicate this condition with a Status Message, nominally of "Query Not Supported".

Queries should be fulfilled in a manner that produces the same result as following these steps:

1. As an optional first step, inspect the Query structure for errors (e.g., a relationship that is not valid for a given Capability Module) and unsupported features (e.g., an unsupported Capability Module or Targeting Expression). If an error or unsupported feature is detected, respond with a Status Message that identifies the error condition.

2. For each record in the identified TAXII Data Collection (the Data Collection name is specified outside of the Query structure), evaluate the Criteria. If the Criteria evaluates to "true" the record should be included in the result set.

Criteria should be evaluated in a manner that produces the same result as following these steps:

1. Create a list of all Child Criteria (Note that Criteria can be a Child of Criteria. For the purposes of this workflow, they are distinguished as the Parent Criteria, which is the Criteria that is evaluated in this workflow, and the Child Criteria, which are immediate descendants of the Parent Criteria) and Child Criterion.
2. For each Child Criteria/Criterion:
   a. If the Child is a Criteria, evaluate the Child Criteria to determine if it is True or False by following this workflow from Step #1.
      (Note: This is recursive. Eventually there will be a Criteria that has only Criterion children.)
   b. If the Child is a Criterion, evaluate the Target against the Test, and apply negation if necessary to determine if the Child Criterion is True or False.
      Note: The authors recognize that this is a non-trivial "exercise left for the reader". However, evaluation of individual Criterion is implementation specific and therefore out of scope for this specification.

8

    c.   If the Child Criteria/Criterion evaluates to True and the Operator is OR, the Parent Criteria evaluates to True.

    d.   If the Child Criteria/Criterion evaluates to True and the Operator is AND, processing continues unless there are no more Child Criteria/Criterion. If there are no more Child Criteria/Criterion, the Parent Criteria evaluates to True.

    e.   If the Child Criteria/Criterion evaluates to False and the Operator is OR, processing continues unless there are no more Child Criteria/Criterion. If there are no more Child Criteria/Criterion, the Parent Criteria evaluates to False.

    f.   If the Child Criteria/Criterion evaluates to False and the Operator is AND, the Parent Criteria evaluates to False.

3   Targeting Expressions

A Targeting Expression is contained by the Target property of a query. Within a Criterion, the Target is used to identify a specific region of a record to which the Test should be applied. This section defines the Targeting Expression syntax used by all TAXII Default Queries. The Targeting Expression syntax, in conjunction with a Targeting Expression Vocabulary, are used to form a Targeting Expression. This section defines one Targeting Vocabulary that Query providers may choose to use. Third parties may define additional vocabularies for use with the Targeting Expression syntax defined by this section.

3.1   Targeting Expression Syntax

All Targeting Expressions use a syntax called Slash Notation. Using the Slash Notation Targeting Expression syntax, a Targeting Expression consists of one or more of Nodes (recall that one or more Nodes make up a Targeting Expression Vocabulary) separated by a forward slash (/). A Node can be one of four things:

1. Node – The name of a Node in the indicated Targeting Expression Vocabulary (This is indicated by the Targeting Expression ID property of a Query). Field Names are case sensitive unless the Targeting Expression Vocabulary defines them to be case insensitive.
2. Field Wildcard – This indicates any Node. Only a single Node is represented. This is indicated by a star (*).
3. Multi-field Wildcard – This indicates any Node or series of Nodes. This is indicated by two stars (**).

3.2   Targeting Expression Vocabularies

A Targeting Expression vocabulary defines which Nodes are permitted in a Targeting Expression, the Node hierarchy, and whether wildcards are permitted. Targeting Expression Vocabularies can range from a list of allowed Nodes to hierarchy of Nodes.

This document defines one Targeting Expression Vocabulary for STIX, which query providers may choose to use (or not). Third parties may define their own Targeting Expression Vocabularies.

3.2.1 STIX Targeting Expression Vocabulary

The Targeting Expression Vocabulary ID that identifies the STIX Targeting Expression Vocabulary is the Content Binding ID for STIX. Recall that the formula for a STIX Content Binding ID is:

$$\texttt{"urn:stix.mitre.org:"} + format + \texttt{":"} + version$$

The set of allowed Nodes within a Targeting Expression using this vocabulary are:

1. Any XML element defined by the version of STIX identified by the *version* portion of the Targeting Expression Vocabulary ID. These Nodes do not have any additional marking (e.g., the 'STIX_Package' element Node name is 'STIX_Package').
2. Any XML attribute defined by the version of STIX identified by the *version* portion of the Targeting Expression Vocabulary ID. These Nodes are prefixed by an at (@) symbol (e.g., the 'version' attribute Node name is '@attribute').

The Node ordering is defined by the version of STIX identified by the *version* portion of the Targeting Expression Vocabulary ID. Specifically, the Node hierarchy follows the following rules:

1. The STIX root element (e.g., STIX_Package) is the root Node and is at the top of the hierarchy.
2. Child elements and attributes of a STIX element are children of that Node
    a. e.g., 'Indicators', an XML element, and 'version', an XML attribute, are both child Nodes of the STIX_Package Node.
    b. The 'Indicators' Node name is 'Indicators'
    c. The 'version' Node name is '@version'
3. The Field Wildcard (*) is permitted.
4. The Multi-field Wildcard (**) is permitted.

Examples:

1. STIX_Package/* targets any element or attribute child of the STIX_Package XML Element
2. STIX_Package/Indicators/** targets any element or attribute descendant of the Indicators XML Element.
3. **/@id targets to any element or attribute named 'id' within the STIX structure.

3.2.2 Third Party Targeting Expression Vocabularies

All Third Party Targeting Expression Vocabularies MUST define the following information:

1. The Targeting Expression Vocabulary ID, which MUST be in URI format.
2. The set of allowed Nodes
3. The hierarchy of allowed nodes
4. The meaning of the Field Wildcard (the Field Wildcard MAY be prohibited)
5. The meaning of the Multi-field Wildcard (the Multi-field Wildcard MAY be prohibited)
6. At least one example Targeting Expression. The example should include a statement as to which record region is targeted by that Targeting Expression.

3.2.3 Example Third Party Targeting Expression Vocabulary

This section provides an example that only permits a single field of "File_Hash". A Third Party might define this vocabulary if they wish to provide a service that permits only queries that look for information on a particular file hash.

Targeting Expression Vocabulary ID: urn:example.com:vocab:filehash
Allowed Nodes: 'File_Hash'
Node Hierarchy: There is no hierarchy, as there is only one level of Node
Field Wildcard: This is prohibited
Multi-field Wildcard: This is prohibited
Examples:

1. File_Hash targets the file hash portion of the record.

4 Capability Modules

This section contains the Capability Modules defined by this document. Third parties may define additional capability modules for use with the TAXII Default Query.

This section defines thee capability modules:

- Core – A common set of relationships that are expected to be implementable across a wide range of systems.
- Regular Expression – Defines the ability to use a regular expression in a Default Query.
- Timestamps – Relationships that can be used to compare timestamps.

4.1 Capability Module: Core

This section defines the Core Capability Module. The Core Capability Module includes a set of relationships that can be expressed in a wide range of database systems.

The Capability Module ID that identifies this capability module is:

$$\texttt{urn:taxii.mitre.org:query:capability:core-1}$$

4.1.1 Relationship: equals

The equals relationship returns true if the target matches the value exactly. If the target merely contains the value (but does not match exactly) the relationship returns false.

Table 4 - Parameters for Core Equals

Parameter Name	Permitted Values	Description
match_type	Only the following values are permitted: • case_sensitive_string	case_sensitive_string indicates that a case sensitive string comparison should be performed.

Parameter Name	Permitted Values	Description
	• case_insensitive_string • number	case_insensitive_string indicates that a case insensitive string comparison should be performed. number indicates that a numeric comparison should be performed. Other match types (e.g., Date/Time) are not permitted for this relationship.
value	Any string is permitted	The string that the target is compared against.

4.1.2   Relationship: not equals

The not equals relationship returns true if the target does not match the value.

Table 5 - Parameters for Core Not Equals

Parameter Name	Permitted Values	Description
match_type	Only the following values are permitted: • case_sensitive_string • case_insensitive_string • number	case_sensitive_string indicates that a case sensitive string comparison should be performed. case_insensitive_string indicates that a case insensitive string comparison should be performed. number indicates that a numeric comparison should be performed. Other match types (e.g., Date/Time) are not permitted for this relationship.
value	Any string is permitted	The string that the target is compared against.

4.1.3   Relationship: greater than

The greater than relationship returns true if the target is numerically greater than the value. This relationship is only valid for numeric comparisons (e.g., it is not valid for string comparisons).

Table 6 - Parameters for Core Greater Than

Parameter Name	Permitted Values	Description
value	Any number is permitted	The number that the target is compared against.

4.1.4   Relationship: greater than or equal

The greater than or equal relationship returns true if the target is numerically greater than or equal to the value. This relationship is only valid for numeric comparisons (e.g., it is not valid for string comparisons).

Table 7 - Parameters for Core Greater Than or Equals

Parameter Name	Permitted Values	Description
value	Any number is permitted	The number that the target is compared against.

4.1.5   Relationship: less than

The less than relationship returns true if the target is numerically less than the value. This relationship is only valid for numeric comparisons (e.g., it is not valid for string comparisons).

Table 8 - Parameters for Core Less Than

Parameter Name	Permitted Values	Description
value	Any number is permitted	The number that the target is compared against.

4.1.6   Relationship: less than or equal

The less than or equal relationship returns true if the target is numerically less than or equal to the value. This relationship is only valid for numeric comparisons (e.g., it is not valid for string comparisons).

Table 9 - Parameters for Core Less Than or Equal

Parameter Name	Permitted Values	Description
value	Any number is permitted	The number that the target is compared against.

4.1.7   Relationship: does not exist

The greater than relationship returns true if the target does not exist.

Table 10 - Parameters for Core Does Not Exist

Parameter Name	Permitted Values	Description
There are not any parameters for this relationship.		

4.1.8   Relationship: exists

The contains relationship returns true if the target exists.

Table 11 - Parameters for Core Exists

Parameter Name	Permitted Values	Description

13

There are not any parameters for this relationship.	

4.1.9   Relationship: begins with

The begins with relationship returns true if the target begins with the value. This relationship is only valid for string comparisons.

Table 12 - Parameters for Core Begins With

Parameter Name	Permitted Values	Description
case_sensitive	Only the following values are permitted: • true • false	If true, a case sensitive comparison should be performed. If false, a case insensitive comparison should be performed. If this field is absent, this parameter should be treated as "true".
value	Any string is permitted	The string that the target is compared against.

4.1.10   Relationship: ends with

The ends with relationship returns true if the target ends with the value. This relationship is only valid for string comparisons.

Table 13 - Parameters for Core Ends With

Parameter Name	Permitted Values	Description
case_sensitive	Only the following values are permitted: • true • false	If true, a case sensitive comparison should be performed. If false, a case insensitive comparison should be performed. If this field is absent, this parameter should be treated as "true".
value	Any string is permitted	The string that the target is compared against.

4.1.11   Relationship: contains

The contains relationship returns true if the target contains the value. This relationship is only valid for string comparisons.

Table 14 - Parameters for Core Contains

Parameter Name	Permitted Values	Description
case_sensitive	Only the following values are permitted: • true • false	If true, a case sensitive comparison should be performed. If false, a case insensitive comparison should be performed. If this field is absent, this parameter should be treated as "true".
value	Any string is permitted	The string that the target is compared against.

4.2   Capability Module: Regular Expression

This section defines the Regular Expression Capability Module. The Regular Expression Capability Module includes a single relationship that is used to perform Regular Expression Matching.

The Capability Module ID that identifies this capability module is:

<div align="center">

`urn:taxii.mitre.org:query:capability:regex-1`

</div>

4.2.1   Relationship: matches

The matches relationship returns true if the target matches the regular expression contained in the value.

<div align="center">

Table 15 - Parameters for Regex Matches

</div>

Parameter Name	Permitted Values	Description
case_sensitive	Only the following values are permitted: • true • false	true indicates that the regular expression should be matched in a case sensitive manner. False indicates that the regular expression should be matched in a case insensitive manner.
value	Regular expressions that conform to the CybOX common subset of regular expression syntax.	The regular expression that the target is compared against. The regular expressions in this field must conform to the regular expression syntax used by CybOX: http://cybox.mitre.org/language/regular_expression_support.pdf.

4.3   Capability Module – Timestamp

The Capability Module ID that identifies this capability module is:

<div align="center">

`urn:taxii.mitre.org:query:capability:timestamp-1`

</div>

This capability module includes relationships that operate on timestamps.

4.3.1   Relationship: equals

The equals relationship returns true if the target and the value indicate the same time and date. This relationship is only valid for timestamp comparisons.

<div align="center">

Table 16 - Parameters for Timestamp Equals

</div>

Parameter Name	Permitted Values	Description
value	Any RFC 3339 conformant timestamp is permitted	The timestamp that the target is compared against.

15

4.3.2 Relationship: greater than

The greater than relationship returns true if the target occurs after the value. This relationship is only valid for timestamp comparisons.

<div align="center">Table 17 - Parameters for Timestamp Greater Than</div>

Parameter Name	Permitted Values	Description
value	Any RFC 3339 conformant timestamp is permitted	The timestamp that the target is compared against.

4.3.3 Relationship: greater than or equals

The greater than or equals relationship returns true if the target occurs after the value or the target and value indicate the same time and date. This relationship is only valid for timestamp comparisons.

<div align="center">Table 18 - Parameters for Timestamp Greater Than or Equals</div>

Parameter Name	Permitted Values	Description
value	Any RFC 3339 conformant timestamp is permitted	The timestamp that the target is compared against.

4.3.4 Relationship: less than

The less than relationship returns true if the target occurs before the value. This relationship is only valid for timestamp comparisons.

<div align="center">Table 19 - Parameters for Timestamp Less Than</div>

Parameter Name	Permitted Values	Description
value	Any RFC 3339 conformant timestamp is permitted	The timestamp that the target is compared against.

4.3.5 Relationship: less than or equals

The less than or equals relationship returns true if the target occurs before the value or the target and value indicate the same time and date. This relationship is only valid for timestamp comparisons.

<div align="center">Table 20 - Parameters for Timestamp Less Than or Equals</div>

Parameter Name	Permitted Values	Description
value	Any RFC 3339 conformant timestamp is permitted	The timestamp that the target is compared against.

5 Examples

5.1 Query Information Structure Example

```
<!-- Query Information example -->
<DefaultQueryInformation xmlns="http://taxii.mitre.org/query/taxii_default_query-1">
    <TargetingExpressionID>urn:taxii.mitre.org:query:vocab:stix-1.0</TargetingExpressionID>
    <TargetingExpressionID>urn:taxii.mitre.org:query:vocab:stix-1.1</TargetingExpressionID>
    <CapabilityModule>urn:taxii.mitre.org:query:capability:core-1</CapabilityModule>
    <CapabilityModule>urn:taxii.mitre.org:query:capability:regex-1</CapabilityModule>
</DefaultQueryInformation>
```

5.2 Query Structure Example - 001

```
<!-- This tests for id attributes that begin with 'EXAMPLE' (case insensitive) -->
<DefaultQuery xmlns="http://taxii.mitre.org/query/taxii_default_query-1"
targeting_expression_id="urn:taxii.mitre.org:query:vocab:stix-1.1">
    <Criteria operator="OR">
        <Criterion negate="false">
            <Target>**/@id</Target><!-- Matches any ID attribute, anywhere -->
            <Test capability_id="urn:taxii.mitre.org:query:capability:core-1" relationship="begins with">
                <Parameter name="case_sensitive">false</Parameter>
                <Parameter name="value">EXAMPLE</Parameter>
            </Test>
        </Criterion>
    </Criteria>
</DefaultQuery>
```

5.3 Query Structure Example – 002

```
<!-- This tests for id attributes that begin with 'example' (case sensitive) and contains 'The quick brown
fox jumped over the very lazy dogs.' -->
<DefaultQuery xmlns="http://taxii.mitre.org/query/taxii_default_query-1"
targeting_expression_id="urn:taxii.mitre.org:query:vocab:stix-1.1">
    <Criteria operator="AND">
        <Criterion negate="false">
            <Target>**/@id</Target><!-- Matches any ID attribute, anywhere -->
            <Test capability_id="urn:taxii.mitre.org:query:capability:core-1" relationship="begins with">
                <Parameter name="case_sensitive">true</Parameter>
                <Parameter name="value">example</Parameter>
            </Test>
        </Criterion>
        <Criterion negate="false">
            <Target>**/Description</Target><!-- Matches any ID attribute, anywhere -->
            <Test capability_id="urn:taxii.mitre.org:query:capability:core-1" relationship="contains">
                <Parameter name="case_sensitive">false</Parameter>
                <Parameter name="value">The quick brown fox jumped over the very lazy dogs.</Parameter>
```

```
        </Test>
      </Criterion>
    </Criteria>
  </DefaultQuery>
```