stix2-validator Documentation

Release 2.0.1

OASIS Open

Jun 10, 2020

Contents:

1	Installation	3
2	Usage2.1As A Script	
3	Options	7
4	Checking STIX Content4.1Mandatory Checks - STIX 2.14.2Optional Checks - STIX 2.1	
5	Contributing5.1Setting up a development environment5.2Code style5.3Testing5.4Adding a dependency	24 24
6	Indices and tables	25

The STIX Validator checks that STIX JSON content conforms to the requirements specified in the latest STIX 2 specifications. In addition to checking conformance with the JSON schemas, the validator checks conformance with requirements that cannot be specified in JSON schema, as well as with established "best practices." This validator is non-normative; in cases of conflict with the STIX specification, the specification takes precedence.

The STIX 2 specification contains two types of requirements: mandatory "MUST" requirements, and recommended "SHOULD" best practice requirements. The validator checks documents against the "MUST" requirements using JSON schemas. Some of these mandatory requirements cannot be implemented in JSON Schema, however, so the validator uses Python functions to check them. The "SHOULD" requirements are all checked by Python functions, and options may be used to ignore some or all of these recommended "best practices."

The STIX Validator uses the stix2-patterns validator to check that Indicator patterns conform to the STIX Patterning language and only reference properties valid for the objects in the pattern.

The validator also color-codes its output to make it easier to tell at a glance whether validation passed.

Installation

Note: The STIX 2 validator requires Python 2.7 or 3.4+.

The easiest way to install the STIX 2 validator is with pip:

```
$ pip install stix2-validator
```

Note that if you instead install it by cloning or downloading the repository, you will need to set up the submodules before you install it:

```
$ git clone https://github.com/oasis-open/cti-stix-validator.git
$ cd cti-stix-validator/
$ git submodule update --init --recursive
$ python setup.py install
```

Usage

2.1 As A Script

The validator comes with a bundled script which you can use to validate a JSON file containing STIX content:

```
$ stix2_validator <stix_file.json>
```

2.2 As A Library

You can also use this library to integrate STIX validation into your own tools. You can validate a JSON file:

```
from stix2validator import validate_file, print_results
results = validate_file("stix_file.json")
print_results(results)
```

You can also validate a JSON string, and check if the input passed validation:

```
from stix2validator import validate_string, print_results
stix_json_string = "..."
results = validate_string(stix_json_string)
if results.is_valid:
    print_results(results)
```

If your STIX is already in a Python dictionary (for example if you have already run json.loads()), use validate_instance() instead:

```
import json
from stix2validator import validate_instance, print_results
```

(continues on next page)

(continued from previous page)

```
stix_json_string = "..."
stix_obj = json.loads(stix_json_string)
results = validate_instance(stix_obj)
if results.is_valid:
    print_results(results)
```

You can pass a ValidationOptions object into validate_file(), validate_string(), or validate_instance() if you want behavior other than the default:

```
from stix2validator import ValidationOptions
options = ValidationOptions(strict=True)
results = validate_string(stix_json_string, options)
```

2.3 STIX 2 Versions

By default the validator will check content against the latest version of the STIX 2 specification. However, older versions can be checked with the version option. For example:

\$ stix2_validator --version=2.0 <stix_file.json>

or in Python:

```
options = ValidationOptions(strict=True, version="2.0")
results = validate_string(stix_json_string, options)
```

2.4 Additional Schemas

The validator uses the STIX 2 JSON schemas as the basis for its validation, but you can also validate with your own additional schemas. This can help if you want to validate STIX content using custom objects, properties, or observables.

To do this use the --schema-dir argument:

\$ stix2_validator --schema-dir /path/to/my/schemas <stix_file.json>

or in Python:

You can see some examples of custom schemas here. Note that if you want to add a custom property to an existing object type, your custom schema only needs to contain that property; the validator's built-in schemas are still checked against and will handle the rest.

chapter $\mathbf{3}$

Options

These are the different options that can be set, whether the validator is used as a command-line script or as a Python library. When using the validator as a library, these options can be passed as parameters to the ValidationOptions constructor.

Script	Library	Description	
FILES	files	A whitespace separated list of STIX files or directories of STIX files to	
		validate.	
-r,recursive		v Recursively descend into input directories.	
-s SCHEMA_DIR,	schema_o	Chustom schema directory. If provided, input will be validated against	
schemas		these schemas in addition to the STIX schemas bundled with this script.	
SCHEMA_DIR			
version	version	The version of the STIX specification to validate against (e.g. "2.0").	
-v,verbose	verbose	Print informational notes and more verbose error messages.	
-q,silent	silent	Silence all output to stdout.	
-d DISABLED,	disable	A comma-separated list of recommended best practice checks to skip.	
disable		By default, no checks are disabled. Example: -disable 202,210	
DISABLED,ignore			
DISABLED			
-e ENABLED,	enabled	I I I I I I I I I I I I I I I I I I I	
enable ENABLED,		If the –disable option is not used, no other checks will be run. By default,	
select ENABLED		all checks are enabled. Example: -enable 218	
strict	strict	Treat warnings as errors and fail validation if any are found.	
strict-types	strict_t	Figures that no custom object types are used, only those defined in the	
		STIX specification.	
strict-properties	strict_P	Ensurethatsno custom properties are used, only those defined in the	
		STIX specification.	
no-cache		e Disable the caching of external source values.	
refresh-cache	refresh	Cars the cache of external source values, then during validation down-	
		loads them again.	
clear-cache		$_{ca}$ Chear the cache of external source values after validation.	
enforce-refs	enforce_	Fanfusres that all SDOs being referenced by SROs are contained within	
		the same bundle.	

For the list of checks that can be used with the "enabled" or "disabled" options, see the Best Practices page.

Checking STIX Content

The validator will always validate input against all of the mandatory "MUST" requirements from the spec. By default it will issue warnings if the input fails any of the "SHOULD" recommendations, but validation will still pass. To turn these "best practice" warnings into errors and cause validation to fail, use the --strict option with the command-line script, or create a ValidationOptions object with strict=True when using the library.

You cannot select which of the "MUST" requirement checks will be performed; all of them will always be performed. However, you may select which of the "SHOULD" checks to perform. Use the codes from the table below to enable or disable these checks. For example, to disable the checks for the report label and tool label vocabularies, use --disable 218,222 or disabled="218,222". All the other checks will still be performed. Conversely, to only check that custom property names adhere to the recommended format but not run any of the other "best practice" checks, use --enable 103 or enabled="103".

Enabling supersedes disabling. Simultaneously enabling and disabling the same check will result in the validator performing that check.

Some checks access Internet resources to determine valid values for certain properties. For instance, the 'mimetype' check accesses the IANA's list of registered MIME types. These web requests are cached to conserve bandwidth, will expire after one week, and are stored in a file called 'cache.sqlite' in the same directory the script is run from. The cache can be refreshed manually with the --refresh-cache or refresh_cache=True, or cleared with --clear-cache or clear_cache=True. This caching can be disabled entirely with --no-cache or no_cache=True.

4.1 Mandatory Checks - STIX 2.1

Nam	e Ensures	Errors/Warnings
times	- timestamps contain sane	<pre>'<property>': '<timestamp>' is not a valid timestamp: <message></message></timestamp></property></pre>
tamp	months, days, hours, min-	' <object>': '<property>': '<timestamp>' is not a valid timestamp: <mes-< td=""></mes-<></timestamp></property></object>
	utes, seconds	sage>
		' <object>': '<extension>': '<property>': '<timestamp>' is not a a valid</timestamp></property></extension></object>
		timestamp: <message></message>
		' <object>': '<property>': '<embedded-property>' is not a valid timestamp:</embedded-property></property></object>
		<message></message>
times	- timestamp properties with a	' <operand_1>' (<operand1_value>) must be <comparison_op></comparison_op></operand1_value></operand_1>
tamp	ccompareison are valid	<pre>'<operand_2>' (<operand2_value)< pre=""></operand2_value)<></operand_2></pre>
ob-	cyber observable timestamp	In object ' <identifier>', '<operand_1>' (<operand1_value>) must be <com-< td=""></com-<></operand1_value></operand_1></identifier>
serv-	•	parison_op> ' <operand_2>' (<operand2_value>)</operand2_value></operand_2>
	tisoustampirementaare valid	
ob-	that marking definitions do	'object_marking_refs' cannot contain any references to this marking defini-
	n adding<u>o</u>níain larinæfkar refer-	tion object (no circular references)
	ences (i.e., they do not refer-	
	ence themselves in the 'ob-	
	ject_marking_refs' property	
gran-		'granular markings' cannot contain any references to this marking definition
u-	do not contain circular	object (no circular references)
	arkfngsnceisculaie_refsthey do	object (no circular references)
1a1_11	not reference themselves	
	in the 'granular_markings'	
	property	' <selector>' is not a valid selector because '<index>' is not a valid index</index></selector>
	selectors in granular mark-	
ing_s	elingor <u>e</u> fyntaxitems which are	' <selector>' is not a valid selector because '<selector_segment>' is not a</selector_segment></selector>
	actually present in the object	list.
		' <selector>' is not a valid selector because '<selector_segment>' is not a</selector_segment></selector>
1		property.
ob-	certain observable object	' <property>' in observable object '<identifier>' can't resolve '<embed-< td=""></embed-<></identifier></property>
serv-	1 1	property>' reference ' <identifier>'</identifier>
able_	ologionate_atefigpencefesbject	' <property>' in observable object '<identifier>' must refer to an object of</identifier></property>
		type ' <type(s)>'</type(s)>
ar-	the 'mime_type' property of	the 'mime_type' property of object ' <identifier>' ('<mime_type>') must be</mime_type></identifier>
ti-	artifact objects comes from	an IANA registered MIME Type of the form 'type/subtype'.
fact_	mithe Treppplate column in the	
L	IANA media type registry	
char-		The 'path_enc' property of object ' <identifier>' ('<path_enc>') must be an</path_enc></identifier>
ac-	ber observable objects come	IANA registered character set.
ter_s	t from the IANA Character	The 'name_enc' property of object ' <identifier>' ('<name_enc>') must be</name_enc></identifier>
	Set list.	IANA registered character set.
lan-	the 'lang' property of SDOs	' <lang>' is not a valid RFC 5646 language code.</lang>
guage	e is a valid RFC 5646 lan-	
	guage code	
soft-	the 'language' property of	The 'languages' property of object ' <identifier>' contains an invalid code</identifier>
ware	langtwage objects is a valid	(' <lang>').</lang>
	ISO 639-2 language code	
pat-	that the syntax of the pattern	' <object>' is not a valid observable type name</object>
terns	of an indicator is valid, and	Custom Observable Object type ' <object>' should start with 'x-' followed</object>
	that objects and properties	by a source unique identifier (like a domain name with dots replaced by
	referenced by the pattern	hyphens), a hyphen and then the name
	are valid. This runs the	Custom Observable Object type ' <object>' should start with 'x-'</object>
4.1. N	latidatory Chiecksr- STIX-2.	
n	//github.com/oasis-open/cti-	Cyber Observable Object custom property ' <property>' should start with</property>
	pattern-validator) to check	'x_' followed by a source unique identifier (like a domain name with dots
	the syntax of the pattern.	replaced by underscores), an underscore and then the name
	• <u>1</u>	Cyber Observable Object sustern property 'correspondy's should start with

4.2 Optional Checks - STIX 2.1

Code	Name	Ensures	Errors/Warnings
1	format-checks	all 1xx checks are run.	
		Specifically:	
			Continued on next page

Chapter 4. Checking STIX Content

	Table 1 – continued	from previous page	
101	Table 1 – continued custom-prefix	from previous page names of custom object types, properties, observ- able objects, observable object properties, and ob- servable object extensions follow the correct format	custom object type ' <object>' should start with 'x- ' followed by a source unique identifier (like a domain name with dots re- placed by hyphens), a hy- phen and then the name. custom property '<prop- erty>' should have a type</prop- </object>
			that starts with 'x_' fol- lowed by a source unique identifier (like a domain name with dots replaced by a hyphen), a hyphen and then the name. Custom Observable Object type ' <observ-< th=""></observ-<>
			able_object>' should start with 'x-' followed by a source unique identifier (like a domain name with dots replaced by hyphens), a hyphen and then the name. Custom Cyber Observ-
			able Object extension type ' <observable-object- extension>' should start with 'x-' followed by a source unique identifier (like a domain with dots replaced by hyphens), a hyphen and then the</observable-object-
			a hyphen and then the name. Cyber Observ- able Object custom property ' <observ- able_object_property>' should start with 'x_' followed by a source</observ-
			unique identifier (like a domain name with dots replaced by hyphens), a hyphen and then the name. Cyber Observable Object custom property ' <prop-< td=""></prop-<>
			erty>' in the <extension> extension should start with 'x_' followed by a source unique (like a domain name with dots replaced by hyphens), a hyphen and then the</extension>
4.2. Optional Checks - S	TIX 2.1		name. Cyber Observable Ob 13 ject custom property ' <property>' in the <ex- tension_property> of the</ex- </property>

Table 1 – continued from previous page

102	Table I – Continued	same as 101 but more le-	austom object type 'sch
102	custom-prefix-lax		custom object type ' <ob-< td=""></ob-<>
		nient; no source identifier	ject>' should start with 'x-
		needed in prefix	' in order to be compatible
			with future versions of the
			STIX 2 specification.
			custom property ' <prop-< td=""></prop-<>
			erty>' should have a type
			that starts with 'x_' in
			order to be compatible
			with future versions of the
			STIX 2 specification.
			Custom Observable
			Object type ' <observ-< td=""></observ-<>
			able_object>' should start
			with 'x-'.
			Custom Observable
			Object extension
			type ' <observable-< td=""></observable-<>
			object_extension>'
			should start with 'x-'.
			Cyber Observable Object
			custom property ' <prop-< td=""></prop-<>
			erty>' should start with
			'X_'.
			Cyber Observable Object
			custom property ' <em-< td=""></em-<>
			bedded_property>' in the
			<pre>conded_property> in the <property> of the <ob-< pre=""></ob-<></property></pre>
			ject> object should start
			with 'x_'.
			Cyber Observable Object
			custom property ' <prop-< td=""></prop-<>
			erty>' in the <extension></extension>
			extension should start
			with 'x_'.
			Cyber Observable Ob-
			ject custom property
			' <property>' in the</property>
			<extension_property></extension_property>
			property of the <exten-< td=""></exten-<>
			sion> extension should
			start with 'x_'.
103	uuid-check	objects use the recom-	Cyber Observable ID
		mended versions of UUID	value <identifier> is not a</identifier>
		(v5 for SCOs, v4 for the	valid UUIDv5 ID.
		rest)	Given ID value <iden-< td=""></iden-<>
			tifier> is not a valid
			UUIDv4 ID.
			Continued on next page
			1 9 -

Tahla	1_	continued	from	previous page	
rable	1 -	continuea	IIOIII	previous page	

Chapter 4. Checking STIX Content

111	open-vocab-format	values of open vocabular- ies follow the correct for- mat	Open vocabulary value ' <value>' should be all lowercase and use hy-</value>
			phens instead of spaces or
			underscores as word sepa-
121	kill-chain-names	bill shain phase name and	rators. kill_chain_name
121	kin-cham-hames	kill-chain-phase name and phase follow the correct	' <chain_name>' should</chain_name>
		format	be all lowercase and use
		Tormat	hyphens instead of spaces
			or underscores as word
			separators.
			phase_name
			' <phase_name>' should</phase_name>
			be all lowercase and use
			hyphens instead of spaces
			or underscores as word
			separators
141	observable-object-keys	observable object keys	' <key_value>' is not a</key_value>
		follow the correct format	good key value. Observ-
			able Objects should use
			non- negative integers for
			their keys.
142	observable-dictionary-	dictionaries in cyber ob-	As a dictionary key,
	keys	servable objects follow	' <key_value>' should be</key_value>
1.42		the correct format	lowercase.
143	malware-analysis-product	malware analysis product	The 'product' property
		names follow the correct	of object ' <identifier>'</identifier>
		format	should be all lowercase
			with words separated by dash.
149	windows-process-	windows-process-ext's	The 'priority' property
1.12	priority-format	'priority' follows the	of object ' <identifier>'</identifier>
	P	correct format	should end in '_CLASS'.
	1	1	Continued on next page

Table 1 – continued from previous page	ae
--	----

150	hash-length	keys in 'hashes'-type	Object ' <identifier>'</identifier>
150	nash-iengui	properties are not too long	has a 'hashes' dictio-
		properties are not too long	nary with a hash of type
			' <hash_type>', which is</hash_type>
			longer than 30 characters.
			Object ' <identifier>' has</identifier>
			an NTFS extension with
			an alternate data stream
			that has a 'hashes' dictio-
			nary with a hash of type
			' <hash_type>', which is</hash_type>
			longer than 30 characters.
			Object ' <identifier>' has</identifier>
			a Windows PE Binary
			File extension with a file
			header hash of ' <hash>',</hash>
			which is longer than 30
			characters.
			Object ' <identifier>' has a</identifier>
			Windows PE Binary File
			extension with an optional
			header that has a hash of
			' <hash>', which is longer</hash>
			than 30 characters.
			Object ' <identifier>' has</identifier>
			a Windows PE Binary
			File extension with a sec-
			tion that has a hash of
			' <hash>', which is longer</hash>
			than 30 characters.
			Object ' <identifier>'</identifier>
			hash a 'hashes' dictio-
			nary with a hash of type
			' <hash_type>', which is</hash_type>
			longer than 30 characters.
2	approved-values	all 2xx checks are run.	
		Specifically:	
201	marking-definition-type	marking definitions use a	Marking definition 'def-
		valid definition_type	inition_type' should
			be one of: <marking-< td=""></marking-<>
			definition-type>.
L			Continued on next page

Table	1 – continued from previous page	

	Table 1 – continu	ed from previous page	
202	relationship-types	relationships are among those defined in the spec- ification	<pre>'<object>' is not a sug- gested relationship source object for the '<relation- ship>' relationship. '<relationship>' is not a suggested relation- ship type for '<object>' objects. '<object>' is not a sug- gested relationship target object for '<object>' ob- jects with the '<relation- ship>' relationship.</relation- </object></object></object></relationship></relation- </object></pre>
203	duplicate-ids	objects in a bundle with duplicate IDs have differ- ent <i>modified</i> timestamps	Duplicate ID ' <iden- tifier>' has identical 'modified' timestamp. If they are different versions of the same object, they should have different 'modified' properties,</iden-
210	all-vocabs	all of the following open vocabulary checks are run	<pre>'<property>' contains a value not in the <vocab_name>- ov vocabulary.</vocab_name></property></pre>
211	attack-motivation	certain property values are from the attack- motivation vocabulary	' <property>' contains a value not in the attack- motivation-ov vocabulary</property>
212	attack-resource-level	certain property values are from the attack-resource- level vocabulary	' <property>' contains a value not in the attack- resource-level-ov vocabu- lary</property>
213	identity-class	certain property values are from the identity-class vo- cabulary	' <property>' contains a value not in the identity- class-ov vocabulary</property>
214	indicator-types	certain property values are from the indicator-types vocabulary	' <property>' contains a value not in the indicator- types-ov vocabulary</property>
215	industry-sector	certain property values are from the industry-sector vocabulary	' <property>' contains a value not in the industry- sector-ov vocabulary</property>
216	malware-types	certain property values are from the malware-types vocabulary	' <property>' contains a value not in the malware- types-ov vocabulary</property>
218	report-types	certain property values are from the report-types vo- cabulary	' <property>' contains a value not in the report- types-ov vocabulary</property>
219	threat-actor-types	certain property values	' <property>' contains a</property>

Table 1 – continued from previous page

220	threat actor role	contain property veluce are	' <property>' contains a</property>
220	threat-actor-role	certain property values are	
		from the threat_actor_role	value not in the threat-
221		vocabulary	actor-role-ov vocabulary
221	threat-actor-sophistication	certain property val-	' <property>' contains a</property>
		ues are from the	value not in the threat-
		threat_actor_sophistication	actor-sophistication-ov
		vocabulary	vocabulary
222	tool-types	certain property values are	' <property>' contains a</property>
		from the tool_types vo-	value not in the tool-
		cabulary	types-ov vocabulary
223	region	certain property values are	' <property>' contains a</property>
		from the region vocabu-	value not in the region-ov
		lary	vocabulary
225	grouping-context	certain property values are	' <property>' contains a</property>
		from the grouping-context	value not in the grouping-
		vocabulary	context-ov vocabulary
226	implementation-	certain property values are	' <property>' con-</property>
	languages	from the implementation-	tains a value not in
		languages vocabulary	the implementation-
		•	languages-ov vocabulary
227	infrastructure-types	certain property values are	' <property>' contains</property>
		from the infrastructure-	a value not in the
		types vocabulary	infrastructure-types-ov
		, I , I , I , I , I , I , I , I , I , I	vocabulary
228	malware-capabilities	certain property values	' <property>' contains a</property>
	1	are from the malware-	value not in the malware-
		capabilities vocabulary	capabilities-ov vocabulary
230	processor-architecture	certain property values	' <property>' contains a</property>
	r	are from the processor-	value not in the processor-
		architecture vocabulary	architecture-ov vocabu-
			lary
231	malware-result	certain property values are	' <property>' contains a</property>
		from the malware-result	value not in the malware-
		vocabulary	result-ov vocabulary
		vocaoulary	Continued on next page

Table 1 – continued from previous page	Table	ontinued from p	revious page
--	-------	-----------------	--------------

241	Table 1 – continued	1 1 0	
241	hash-algo	certain property values are	Object ' <identifier>'</identifier>
		from the hash-algo vocab-	has a 'hashes' dictio-
		ulary	nary with a hash of type
			' <hash_type>', which</hash_type>
			is not a value in the
			hash-algorithm-ov vocab-
			ulary nor a custom value
			prepended with 'x_'.
			Object ' <identifier>'</identifier>
			has an NTFS extension
			with an alternate data
			stream that has a 'hashes'
			dictionary with a hash
			of type ' <hash_type>',</hash_type>
			which is not a value in
			the hash- algorithm-ov
			vocabulary nor a custom
			value prepended with
			'x_'.
			Object ' <identifier>'</identifier>
			has a Windows PE Bi-
			nary File extension with
			a file header hash of
			' <hash_type>', which is</hash_type>
			not a value in the hash-
			algorithm- vocabulary nor
			a custom value prepended
			with 'x_'.
			Object ' <identifier>' has a</identifier>
			Windows PE Binary File
			extension with an optional
			header that has a hash
			of ' <hash_type>', which</hash_type>
			is not a value in the
			hash-algorithm-ov vocab-
			ulary nor a custom value
			-
			prepended with 'x_'.
			Object ' <identifier>' has</identifier>
			a Windows PE Binary
			File extension with a
			section that has a hash
			of ' <hash_type>', which</hash_type>
			is not a value in the
			hash-algorithm-ov vocab-
			ulary nor a custom value
			prepended with 'x_'.
L			Continued on next page

Table 1 – continued from previous page

		ed from previous page	
243	windows-pebinary-type	certain property values are from the windows- pebinary-type vocabulary	Object ' <identifier>' has a Windows PE Binary File extension with a 'pe_type' of '<pe_type>', which is not a value in the windows-pebinary-type-</pe_type></identifier>
244	account-type	certain property values are from the account-type vo- cabulary	ov vocabulary. Object ' <identifier>'is a User Account Object with an 'account_type' of '<ac- count_type>', which is not a value in the account- ture on vacabulary.</ac- </identifier>
245	indicator-pattern-types	certain property values are from the pattern-type vo- cabulary	type-ov vocabulary. ' <property>' contains a value not in the pattern- type-ov vocabulary</property>
270	all-external-sources	all of the following exter- nal source checks are run	
271	mime-type	file.mime_type is a valid IANA MIME type	The 'mime_type' property of object ' <identifier>' ('<mime_type>') should be an IANA registered MIME Type of the form 'type/subtype'.</mime_type></identifier>
272	protocols	certain property values are valid IANA Service and Protocol names	The 'protocols' property of object ' <identifier>' contains a value ('<pro- tocol>') not in IANA Service Name and Transport Protocol Port Number Registry.</pro- </identifier>
273	ipfix	certain property values are valid IANA IP Flow In- formation Export (IPFIX) Entities	The 'ipfix' property of object ' <identifier>' con- tains a key ('<ipfix>') not in IANA IP Flow Informa- tion Export (IPFIX) Enti- ties Registry.</ipfix></identifier>
274	http-request-headers	certain property values are valid HTTP request header names	The 'request_header' property of object ' <iden- tifier>' contains an invalid HTTP header ('<http_request_header>').</http_request_header></iden-
275	socket-options	certain property values are valid socket options	The 'options' property of object ' <identi- fier>' contains a key ('<option>') that is not a valid socket option (SOIICMPIICMP6IIPIIPV6I MCASTITCPIIRLMP)_*. Continued on next page</option></identi-

Table 1 – continued from previous page

The 'document_info_dict' property of object ' <iden- tifier>' contains a key ('<key>') that is not a</key></iden-
tifier>' contains a key
•
(' <key>') that is not a</key>
valid PDF Document In-
formation Dictionary key.
Location 'country' should
be a valid ISO 3166-1
ALPHA-2 Code.
The Network Traffic ob-
ject ' <identifier>' should</identifier>
contain both the 'src_port'
and 'dst_port' properties.
External reference ' <src>'</src>
has a URL but no hash.
Both the name and
description properties
SHOULD be present.
Included property ' <prop-< td=""></prop-<>
erty>' is deprecated
within the indicated spec
version.
filt A J J C a H L C S I C V

Table 1 – continued from previous page

Contributing

We're thrilled that you're interested in contributing to the stix2-validator! Here are some things you should know:

- contribution-guide.org has great ideas for contributing to any open-source project (not just this one).
- All contributors must sign a Contributor License Agreement. See CONTRIBUTING.md in the project repository for specifics.
- If you are planning to implement a major feature (vs. fixing a bug), please discuss with a project maintainer first to ensure you aren't duplicating the work of someone else, and that the feature is likely to be accepted.

Now, let's get started!

5.1 Setting up a development environment

We recommend using a virtualenv.

1. Clone the repository. If you're planning to make pull request, you should fork the repository on GitHub and clone your fork instead of the main repo:

\$ git clone https://github.com/yourusername/cti-stix-validator.git

2. Install develoment-related dependencies and set up git submodules:

```
$ cd cti-stix-validator
$ pip install -r requirements.txt
$ git submodule update --init --recursive
```

3. Install pre-commit git hooks:

\$ pre-commit install

At this point you should be able to make changes to the code.

5.2 Code style

All code should follow PEP 8. We allow for line lengths up to 160 characters, but any lines over 80 characters should be the exception rather than the rule. PEP 8 conformance will be tested automatically by Tox and Travis-CI (see below).

5.3 Testing

Note: All of the tools mentioned in this section are installed when you run pip install -r requirements. txt.

This project uses pytest for testing. We encourage the use of test-driven development (TDD), where you write (failing) tests that demonstrate a bug or proposed new feature before writing code that fixes the bug or implements the features. Any code contributions should come with new or updated tests.

To run the tests in your current Python environment, use the pytest command from the root project directory:

\$ pytest

This should show all of the tests that ran, along with their status.

You can run a specific test file by passing it on the command line:

\$ pytest stix2validator/test/test_<xxx>.py

To ensure that the test you wrote is running, you can deliberately add an assert False statement at the beginning of the test. This is another benefit of TDD, since you should be able to see the test failing (and ensure it's being run) before making it pass.

tox allows you to test a package across multiple versions of Python. Setting up multiple Python environments is beyond the scope of this guide, but feel free to ask for help setting them up. Tox should be run from the root directory of the project:

\$ tox

We aim for high test coverage, using the coverage.py library. Though it's not an absolute requirement to maintain 100% coverage, all code contributions must be accompanied by tests. To run coverage and look for untested lines of code, run:

```
$ pytest --cov=stix2validator
$ coverage html
```

then look at the resulting report in htmlcov/index.html.

All commits pushed to the master branch or submitted as a pull request are tested with Travis-CI automatically.

5.4 Adding a dependency

One of the pre-commit hooks we use in our develoment environment enforces a consistent ordering to imports. If you need to add a new library as a dependency please add it to the *known_third_party* section of *.isort.cfg* to make sure the import is sorted correctly.

Indices and tables

- genindex
- modindex
- search