CybOXTM Version 2.1.1 Part 14: Code Object

Working Draft 01

16 November 2015

Technical Committee:

[OASIS Cyber Threat Intelligence (CTI) TC](https://www.oasis-open.org/committees/cti)

Chair:

Richard Struse ([Richard.Struse@HQ.DHS.GOV](mailto:Richard.Struse@HQ.DHS.GOV)), [DHS Office of Cybersecurity and Communications (CS&C)](http://www.dhs.gov/office-cybersecurity-and-communications)

Editors:

Desiree Beck ([dbeck@mitre.org](mailto:ikirillov@mitre.org)), [MITRE Corporation](http://www.mitre.org/)

Trey Darley ([trey@soltra.com](mailto:trey@soltra.com)), [Soltra](http://www.soltra.com/)

Ivan Kirillov ([ikirillov@mitre.org](mailto:ikirillov@mitre.org)), [MITRE Corporation](http://www.mitre.org/)

Rich Piazza ([rpiazza@mitre.org](mailto:ikirillov@mitre.org)), [MITRE Corporation](http://www.mitre.org/)

Additional artifacts:

This prose specification is one component of a Work Product which consists of:

* *CybOXTM Version 2.1.1 Part 1: Overview*. [URI]
* *CybOXTM Version 2.1.1 Part 2: Common*. [URI]
* *CybOXTM Version 2.1.1 Part 3: Core*. [URI]
* *CybOXTM Version 2.1.1 Part 4: Default Extensions*. [URI]
* *CybOXTM Version 2.1.1 Part 5: Vocabularies*. [URI]
* *CybOXTM Version 2.1.1 Part 6: UML Model*. [URI]
* *CybOXTM Version 2.1.1 Part 7: API Object*. [URI]
* *CybOXTM Version 2.1.1 Part 8: ARP Cache Object*. [URI]
* *CybOXTM Version 2.1.1 Part 9: AS Object*. [URI]
* *CybOXTM Version 2.1.1 Part 10: Account Object*. [URI]
* *CybOXTM Version 2.1.1 Part 11: Address Object*. [URI]
* *CybOXTM Version 2.1.1 Part 12: Archive File Object*. [URI]
* *CybOXTM Version 2.1.1 Part 13: Artifact Object*. [URI]
* *CybOXTM Version 2.1.1 Part 14: Code Object*. (this document)
* *CybOXTM Version 2.1.1 Part 15: Custom Object*. [URI]
* *CybOXTM Version 2.1.1 Part 16: DNS Cache Object*. [URI]
* *CybOXTM Version 2.1.1 Part 17: DNS Query Object*. [URI]
* *CybOXTM Version 2.1.1 Part 18: DNS Record Object*. [URI]
* *CybOXTM Version 2.1.1 Part 19: Device Object*. [URI]
* *CybOXTM Version 2.1.1 Part 20: Disk Object*. [URI]
* *CybOXTM Version 2.1.1 Part 21: Disk Partition Object*. [URI]
* *CybOXTM Version 2.1.1 Part 22: Domain Name Object*. [URI]
* *CybOXTM Version 2.1.1 Part 23: Email Message Object*. [URI]
* *CybOXTM Version 2.1.1 Part 24: File Object*. [URI]
* *CybOXTM Version 2.1.1 Part 25: GUI Dialogbox Object*. [URI]
* *CybOXTM Version 2.1.1 Part 26: GUI Object*. [URI]
* *CybOXTM Version 2.1.1 Part 27: GUI Window Object*. [URI]
* *CybOXTM Version 2.1.1 Part 28: HTTP Session Object*. [URI]
* *CybOXTM Version 2.1.1 Part 29: Hostname Session Object*. [URI]
* *CybOXTM Version 2.1.1 Part 30: Image File Object*. [URI]
* *CybOXTM Version 2.1.1 Part 31: Library File Object*. [URI]
* *CybOXTM Version 2.1.1 Part 32: Link Object*. [URI]
* *CybOXTM Version 2.1.1 Part 33: Linux Package Object*. [URI]
* *CybOXTM Version 2.1.1 Part 34: Memory Object*. [URI]
* *CybOXTM Version 2.1.1 Part 35: Mutex Object*. [URI]
* *CybOXTM Version 2.1.1 Part 36: Network Connection Object*. [URI]
* *CybOXTM Version 2.1.1 Part 37: Network Flow Object*. [URI]
* *CybOXTM Version 2.1.1 Part 38: Network Packet Object*. [URI]
* *CybOXTM Version 2.1.1 Part 39: Network Route Entry Object*. [URI]
* *CybOXTM Version 2.1.1 Part 40: Network Route Object*. [URI]
* *CybOXTM Version 2.1.1 Part 41: Network Socket Object*. [URI]
* *CybOXTM Version 2.1.1 Part 42: Network Subnet Object*. [URI]
* *CybOXTM Version 2.1.1 Part 43: PDF File Object*. [URI]
* *CybOXTM Version 2.1.1 Part 44: Pipe Object*. [URI]
* *CybOXTM Version 2.1.1 Part 45: Port Object*. [URI]
* *CybOXTM Version 2.1.1 Part 46: Process Object*. [URI]
* *CybOXTM Version 2.1.1 Part 47: Product Object*. [URI]
* *CybOXTM Version 2.1.1 Part 48: SMS Message Object*. [URI]
* *CybOXTM Version 2.1.1 Part 49: Semaphore Object*. [URI]
* *CybOXTM Version 2.1.1 Part 50: Socket Address Object*. [URI]
* *CybOXTM Version 2.1.1 Part 51: System Object*. [URI]
* *CybOXTM Version 2.1.1 Part 52: URI Object*. [URI]
* *CybOXTM Version 2.1.1 Part 53: URL History Object*. [URI]
* *CybOXTM Version 2.1.1 Part 54: Unix File Object*. [URI]
* *CybOXTM Version 2.1.1 Part 55: Unix Network Route Entry Object*. [URI]
* *CybOXTM Version 2.1.1 Part 56: Unix Pipe Object*. [URI]
* *CybOXTM Version 2.1.1 Part 57: Unix Process Object*. [URI]
* *CybOXTM Version 2.1.1 Part 58: Unix User Account Object*. [URI]
* *CybOXTM Version 2.1.1 Part 59: Unix Volume Object*. [URI]
* *CybOXTM Version 2.1.1 Part 60: Unix Account Object*. [URI]
* *CybOXTM Version 2.1.1 Part 61: User Session Object*. [URI]
* *CybOXTM Version 2.1.1 Part 62: Volume Object*. [URI]
* *CybOXTM Version 2.1.1 Part 63: Whois Object*. [URI]
* *CybOXTM Version 2.1.1 Part 64: Win Computer Account Object*. [URI]
* *CybOXTM Version 2.1.1 Part 65: Win Critical Section Object*. [URI]
* *CybOXTM Version 2.1.1 Part 66: Win Driver Object*. [URI]
* *CybOXTM Version 2.1.1 Part 67: Win Event Log Object*. [URI]
* *CybOXTM Version 2.1.1 Part 68: Win Event Object*. [URI]
* *CybOXTM Version 2.1.1 Part 69: Win Executable File Object*. [URI]
* *CybOXTM Version 2.1.1 Part 70: Win File Object*. [URI]
* *CybOXTM Version 2.1.1 Part 71: Win Filemapping Object*. [URI]
* *CybOXTM Version 2.1.1 Part 72: Win Handle Object*. [URI]
* *CybOXTM Version 2.1.1 Part 73: Win Hook Object*. [URI]
* *CybOXTM Version 2.1.1 Part 74: Win Kernel Hook Object*. [URI]
* *CybOXTM Version 2.1.1 Part 75: Win Kernel Object*. [URI]
* *CybOXTM Version 2.1.1 Part 76: Win Mailslot Object*. [URI]
* *CybOXTM Version 2.1.1 Part 77: Win Memory Page Region Object*. [URI]
* *CybOXTM Version 2.1.1 Part 78: Win Mutex Object*. [URI]
* *CybOXTM Version 2.1.1 Part 79: Win Network Route Entry Object*. [URI]
* *CybOXTM Version 2.1.1 Part 80: Win Network Share Object*. [URI]
* *CybOXTM Version 2.1.1 Part 81: Win Pipe Object*. [URI]
* *CybOXTM Version 2.1.1 Part 82: Win Prefetch Object*. [URI]
* *CybOXTM Version 2.1.1 Part 83: Win Process Object*. [URI]
* *CybOXTM Version 2.1.1 Part 84: Win Registry Key Object*. [URI]
* *CybOXTM Version 2.1.1 Part 85: Win Semaphore Object*. [URI]
* *CybOXTM Version 2.1.1 Part 86: Win Service Object*. [URI]
* *CybOXTM Version 2.1.1 Part 87: Win System Object*. [URI]
* *CybOXTM Version 2.1.1 Part 88: Win System Restore Object*. [URI]
* *CybOXTM Version 2.1.1 Part 89: Win Task Object*. [URI]
* *CybOXTM Version 2.1.1 Part 90: Win Thread Object*. [URI]
* *CybOXTM Version 2.1.1 Part 91: Win User Account Object*. [URI]
* *CybOXTM Version 2.1.1 Part 92: Win Volume Object*. [URI]
* *CybOXTM Version 2.1.1 Part 93: Win Waitable Timer Object*. [URI]
* *CybOXTM Version 2.1.1 Part 94: X509 Certificate Object*. [URI]

Related work:

This specification is related to:

* *STIXTM Version 1.2.1 (placeholder)*

Abstract:

The Cyber Observable Expression (CybOX) is a standardized language for encoding and communicating high-fidelity information about cyber observables, whether dynamic events or stateful measures that are observable in the operational cyber domain. By specifying a common structured schematic mechanism for these cyber observables, the intent is to enable the potential for detailed automatable sharing, mapping, detection and analysis heuristics. This specification document defines the Code Object data model, which is one of the Object data models for CybOX content.

Status:

This [Working Draft](https://www.oasis-open.org/policies-guidelines/tc-process#dWorkingDraft) (WD) has been produced by one or more TC Members; it has not yet been voted on by the TC or [approved](https://www.oasis-open.org/policies-guidelines/tc-process#committeeDraft) as a Committee Draft (Committee Specification Draft or a Committee Note Draft). The OASIS document [Approval Process](https://www.oasis-open.org/policies-guidelines/tc-process#standApprovProcess) begins officially with a TC vote to approve a WD as a Committee Draft. A TC may approve a Working Draft, revise it, and re-approve it any number of times as a Committee Draft.

URI patterns:

Initial publication URI:  
http://docs.oasis-open.org/cti/stix/v1.2.1/csd01/part1-overview/stix-v1.2.1-csd01-part1-overview.docx

Permanent “Latest version” URI:  
http://docs.oasis-open.org/cti/stix/v1.2.1/stix-v1.2.1-part1-overview.docx

(Managed by OASIS TC Administration; please don’t modify.)

Copyright © OASIS Open 2015. All Rights Reserved.

All capitalized terms in the following text have the meanings assigned to them in the OASIS Intellectual Property Rights Policy (the "OASIS IPR Policy"). The full [Policy](https://www.oasis-open.org/policies-guidelines/ipr) may be found at the OASIS website.

This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published, and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this section are included on all such copies and derivative works. However, this document itself may not be modified in any way, including by removing the copyright notice or references to OASIS, except as needed for the purpose of developing any document or deliverable produced by an OASIS Technical Committee (in which case the rules applicable to copyrights, as set forth in the OASIS IPR Policy, must be followed) or as required to translate it into languages other than English.

The limited permissions granted above are perpetual and will not be revoked by OASIS or its successors or assigns.

This document and the information contained herein is provided on an "AS IS" basis and OASIS DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY OWNERSHIP RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Portions copyright © United States Government 2012-2015.  All Rights Reserved.  
  
STIX™, TAXII™, AND CybOX™ (STANDARD OR STANDARDS) AND THEIR COMPONENT PARTS ARE PROVIDED “AS IS” WITHOUT ANY WARRANTY OF ANY KIND, EITHER EXPRESSED, IMPLIED, OR STATUTORY, INCLUDING, BUT NOT LIMITED TO, ANY WARRANTY THAT THESE STANDARDS OR ANY OF THEIR COMPONENT PARTS WILL CONFORM TO SPECIFICATIONS, ANY IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, OR FREEDOM FROM INFRINGEMENT, ANY WARRANTY THAT THE STANDARDS OR THEIR COMPONENT PARTS WILL BE ERROR FREE, OR ANY WARRANTY THAT THE DOCUMENTATION, IF PROVIDED, WILL CONFORM TO THE STANDARDS OR THEIR COMPONENT PARTS. IN NO EVENT SHALL THE UNITED STATES GOVERNMENT OR ITS CONTRACTORS OR SUBCONTRACTORS BE LIABLE FOR ANY DAMAGES, INCLUDING, BUT NOT LIMITED TO, DIRECT, INDIRECT, SPECIAL OR CONSEQUENTIAL DAMAGES, ARISING OUT OF, RESULTING FROM, OR IN ANY WAY CONNECTED WITH THESE STANDARDS OR THEIR COMPONENT PARTS OR ANY PROVIDED DOCUMENTATION, WHETHER OR NOT BASED UPON WARRANTY, CONTRACT, TORT, OR OTHERWISE, WHETHER OR NOT INJURY WAS SUSTAINED BY PERSONS OR PROPERTY OR OTHERWISE, AND WHETHER OR NOT LOSS WAS SUSTAINED FROM, OR AROSE OUT OF THE RESULTS OF, OR USE OF, THE STANDARDS, THEIR COMPONENT PARTS, AND ANY PROVIDED DOCUMENTATION. THE UNITED STATES GOVERNMENT DISCLAIMS ALL WARRANTIES AND LIABILITIES REGARDING THE STANDARDS OR THEIR COMPONENT PARTS ATTRIBUTABLE TO ANY THIRD PARTY, IF PRESENT IN THE STANDARDS OR THEIR COMPONENT PARTS AND DISTRIBUTES IT OR THEM “AS IS.”

Table of Contents

[1 Introduction 6](#_Toc434322109)

[1.1 CybOXTM Specification Documents 6](#_Toc434322110)

[1.2 Document Conventions 6](#_Toc434322111)

[1.2.1 Fonts 6](#_Toc434322112)

[1.2.2 UML Package References 7](#_Toc434322113)

[1.2.3 UML Diagrams 7](#_Toc434322114)

[1.2.3.1 Class Properties 7](#_Toc434322115)

[1.2.3.2 Diagram Icons and Arrow Types 7](#_Toc434322116)

[1.2.4 Property Table Notation 8](#_Toc434322117)

[1.2.5 Property and Class Descriptions 8](#_Toc434322118)

[1.3 Terminology 9](#_Toc434322119)

[1.4 Normative References 9](#_Toc434322120)

[2 Background Information 10](#_Toc434322121)

[2.1 Cyber Observables 10](#_Toc434322122)

[2.2 Objects 10](#_Toc434322123)

[3 Data Model 11](#_Toc434322124)

[3.1 CodeObjectType Class 11](#_Toc434322125)

[3.2 CodeTypeType Class 13](#_Toc434322126)

[3.3 CodeSegmentXORType Class 13](#_Toc434322127)

[3.4 CodePurposeType Class 15](#_Toc434322128)

[3.5 CodeLanguageType Class 15](#_Toc434322129)

[3.6 ProcessorTypeType Class 16](#_Toc434322130)

[3.7 TargetedPlatformsType Class 16](#_Toc434322131)

[3.8 CodeTypeEnum Enumeration 16](#_Toc434322132)

[3.9 CodePurposeEnum Enumeration 17](#_Toc434322133)

[3.10 CodeLanguageEnum Enumeration 17](#_Toc434322134)

[3.11 ProcessorTypeEnum Enumeration 19](#_Toc434322135)

[4 Conformance 21](#_Toc434322136)

[Acknowledgments 22](#_Toc434322137)

[Revision History 23](#_Toc434322138)

# Introduction

[All text is normative unless otherwise labeled]

The Cyber Observable Expression (CybOXTM) provides a common structure for representing cyber observables across and among the operational areas of enterprise cyber security. CybOX improves the consistency, efficiency, and interoperability of deployed tools and processes, and it increases overall situational awareness by enabling the potential for detailed automatable sharing, mapping, detection, and analysis heuristics.

This document serves as the specification for the CybOX Code Object Version 2.1.1 data model, which is one of eighty-eight CybOX Object data models.

In Section **1.1** we discuss additional specification documents, in Section **1.2** we provide document conventions in, and in Section **1.3** we provide terminology. References are given in Section **1.4**. In Section **2**, we give background information necessary to fully understand the Code Object data model. We present the Code Object data model specification details in Section **3** and conformance information in Section **4**.

## CybOXTM Specification Documents

The CybOX specification consists of a formal UML model and a set of textual specification documents that explain the UML model. Specification documents have been written for each of the key individual data models that compose the full CybOX UML model.

CybOX has a modular design comprising two fundamental data models and a collection of Object data models. The fundamental data models – CybOX Core and CybOX Common – provide essential CybOX structure and functionality. The CybOX Objects, defined in individual data models, are precise characterizations of particular types of observable cyber entities (e.g., HTTP session, Windows registry key, DNS query).

Use of the CybOX Core and Common data models is required; however, use of the CybOX Object data models is purely optional: users select and use only those Objects and corresponding data models that are needed. Importing the entire CybOX suite of data models is not necessary.

The [*CybOX Version 2.1.1 Part 1: Overview*](#AdditionalArtifacts) document provides a comprehensive overview of the full set of CybOX data models, which in addition to the Core, Common, and numerous Object data models, includes various extension data models and a vocabularies data model, which contains a set of default controlled vocabularies. [*CybOX Version 2.1.1 Part 1: Overview*](#AdditionalArtifacts) also summarizes the relationship of CybOX to other languages, and outlines general CybOX data model conventions.

## Document Conventions

The following conventions are used in this document.

### Fonts

The following font and font style conventions are used in the document:

* Capitalization is used for CybOX high level concepts, which are defined in [*CybOX Version 2.1.1 Part 1: Overview*](#AdditionalArtifacts).

Examples: Action, Object, Event, Property

* The Courier New font is used for writing UML objects.

Examples: ActionType, cyboxCommon:BaseObjectPropertyType

Note that all high level concepts have a corresponding UML object. For example, the Action high level concept is associated with a UML class named, ActionType.

* The ‘*italic’* font (withsingle quotes) is used for noting actual, explicit values for CybOX Language properties. The *italic* font (without quotes) is used for noting example values.

Example:  *‘HashNameVocab-1.0,’ high, medium, low*

### UML Package References

Each CybOX data model is captured in a different UML package (e.g., Core package) where the packages together compose the full CybOX UML model. To refer to a particular class of a specific package, we use the format package\_prefix:class, where package\_prefix corresponds to the appropriate UML package. The [*CybOX Version 2.1.1 Part 1: Overview*](#AdditionalArtifacts) document contains the full list of CybOX packages, along with the associated prefix notations, descriptions, and examples.

Note that in this specification document, we do not explicitly specify the package prefix for any classes that originate from the Code Object data model.

### UML Diagrams

This specification makes use of UML diagrams to visually depict relationships between CybOX Language constructs. Note that the diagrams have been extracted directly from the full UML model for CybOX; they have not been constructed purely for inclusion in the specification documents.  Typically, diagrams are included for the primary class of a data model, and for any other class where the visualization of its relationships between other classes would be useful.  This implies that there will be very few diagrams for classes whose only properties are either a data type or a class from the CybOX Common data model.  Other diagrams that are included correspond to classes that specialize a superclass and abstract or generalized classes that are extended by one or more subclasses.

In UML diagrams, classes are often presented with their attributes elided, to avoid clutter. The fully described class can usually be found in a related diagram. A class presented with an empty section at the bottom of the icon indicates that there are no attributes other than those that are visualized using associations.

#### Class Properties

Generally, a class property can be shown in a UML diagram as either an attribute or an association (i.e., the distinction between attributes and associations is somewhat subjective). In order to make the size of UML diagrams in the specifications manageable, we have chosen to capture most properties as attributes and to capture only higher level properties as associations, especially in the main top-level component diagrams. In particular, we will always capture properties of UML data types as attributes. For example, properties of a class that are identifiers, titles, and timestamps will be represented as attributes.

#### Diagram Icons and Arrow Types

Diagram icons are used in a UML diagram to indicate whether a shape is a class, enumeration, or a data type, and decorative icons are used to indicate whether an element is an attribute of a class or an enumeration literal. In addition, two different arrow styles indicate either a directed association relationship (regular arrowhead) or a generalization relationship (triangle-shaped arrowhead). The icons and arrow styles we use are shown and described in **Table 1‑1**.

Table ‑. UML diagram icons

|  |  |
| --- | --- |
| **Icon** | **Description** |
|  | This diagram icon indicates a class. If the name is in italics, it is an abstract class. |
|  | This diagram icon indicates an enumeration. |
|  | This diagram icon indicates a data type. |
|  | This decorator icon indicates an attribute of a class. The green circle means its visibility is public. If the circle is red or yellow, it means its visibility is private or protected. |
|  | This decorator icon indicates an enumeration literal. |
|  | This arrow type indicates a directed association relationship. |
|  | This arrow type indicates a generalization relationship. |

### Property Table Notation

Throughout Section **3** tables are used to describe the properties of each data model class. Each property table consists of a column of names to identify the property, a type column to reflect the datatype of the property, a multiplicity column to reflect the allowed number of occurrences of the property, and a description column that describes the property. Package prefixes are provided for classes outside of the Code Object data model (see Section **1.2.2**).

Note that if a class is a specialization of a superclass, only the properties that constitute the specialization are shown in the property table (i.e., properties of the superclass will not be shown). However, details of the superclass may be shown in the UML diagram.

### Property and Class Descriptions

Each class and property defined in CybOX is described using the format, “The X property verbY.” For example, in the specification for the CybOX Core data model, we write, “The id property specifies a globally unique identifier for the Action.” In fact, the verb “specifies” could have been replaced by any number of alternatives: “defines,” “describes,” “contains,” “references,” etc.

However, we thought that using a wide variety of verb phrases might confuse a reader of a specification document because the meaning of each verb could be interpreted slightly differently. On the other hand, we didn’t want to use a single, generic verb, such as “describes,” because although the different verb choices may or may not be meaningful from an implementation standpoint, a distinction could be useful to those interested in the modeling aspect of CybOX.

Consequently, we have chosen to use the three verbs, defined as follows, in class and property descriptions:

|  |  |
| --- | --- |
| **Verb** | **CybOX Definition** |
| captures | Used to record and preserve information without implying anything about the structure of a class or property. Often used for properties that encompass general content. This is the least precise of the three verbs. |
|  | *Examples*:  The Observable\_Source property characterizes the source of the Observable information. Examples of details captured include identitifying characteristics, time-related attributes, and a list of the tools used to collect the information.  The Description property captures a textual description of the Action. |
| characterizes | Describes the distinctive nature or features of a class or property. Often used to describe classes and properties that themselves comprise one or more other properties. |
|  | *Examples*:  The Action property characterizes a cyber observable Action.  The Obfuscation\_Technique property characterizes a technique an attacker could potentially leverage to obfuscate the Observable. |
| specifies | Used to clearly and precisely identify particular instances or values associated with a property. Often used for properties that are defined by a controlled vocabulary or enumeration; typically used for properties that take on only a single value. |
|  | *Example*:  The cybox\_major\_version property specifies the major version of the CybOX language used for the set of Observables. |

## Terminology

The key words “MUST”, “MUST NOT”, “REQUIRED”, “SHALL”, “SHALL NOT”, “SHOULD”, “SHOULD NOT”, “RECOMMENDED”, “MAY”, and “OPTIONAL” in this document are to be interpreted as described in [**[RFC2119]**](#rfc2119).

## Normative References

[RFC2119] Bradner, S., “Key words for use in RFCs to Indicate Requirement Levels”, BCP 14, RFC 2119, March 1997. <http://www.ietf.org/rfc/rfc2119.txt>.

# Background Information

In this section, we provide high level information about the Code Object data model that is necessary to fully understand the specification details given in Section **3**.

## Cyber Observables

A cyber observable is a dynamic event or a stateful property that occurs, or may occur, in the operational cyber domain. Examples of stateful properties include the value of a registry key, the MD5 hash of a file, and an IP address. Examples of events include the deletion of a file, the receipt of an HTTP GET request, and the creation of a remote thread.

A cyber observable is different than a cyber indicator. A cyber observable is a statement of fact, capturing what was observed or could be observed in the cyber operational domain. Cyber indicators are cyber observable patterns, such as a registry key value associated with a known bad actor or a spoofed email address used on a particular date.

## Objects

Objects in CybOX are individual data models for characterizing a particular cyber entity, such as a Windows registry key, or an Email Message. Accordingly, each release of the CybOX language includes a particular set of Objects that are part of the release. The data model for each of these Objects is defined by its own specification that describes the context-specific classes and properties that compose the Object.

# Data Model

## CodeObjectType Class

The CodeObjectType class is intended to characterize a body of computer code. The UML diagram corresponding to the CodeObjectType class is shown in **Figure 3‑1**.

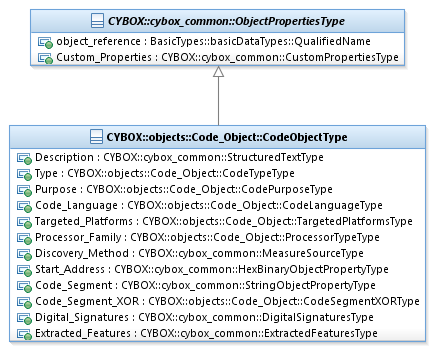


Figure ‑. UML diagram of the CodeObjectType class

The property table of the CodeObjectType class is given in **Table 3‑1**.

Table ‑. Properties of the CodeObjectType class

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Type** | **Multiplicity** | **Description** |
| **Description** | cyboxCommon:  StructuredTextType | 0..1 | The Description property captures a technical description of the CodeObject. Any length is permitted. Optional formatting is supported via the structuring\_format property of the StructuredTextType class. |
| **Type** | CodeObj:CodeTypeType | 0..1 | The Type property is intended to provide a way of specifying the type of code being characterized. |
| **Purpose** | CodeObj:CodePurposeType | 0..1 | The Purpose property is intended to provide a way of specifying the purpose or flavor of code being characterized. |
| **Code\_Language** | CodeObj:CodeLanguageType | 0..1 | The Code\_Language property refers to the code language used in the code characterized in this field. |
| **Targeted\_Platforms** | CodeObj:  TargetedPlatformsType | 0..1 | The Targeted\_Platforms property specifies a list platforms that this code is targeted for. |
| **Processor\_Family** | CodeObj:ProcessorTypeType | 0..\* | The Processor\_Family property specifies the class of processor that the code snippet is targeting. This field may be specified multiple times for code snippets that are applicable across multiple processor families. |
| **Discovery\_Method** | cyboxCommon:  MeasureSourceType | 0..1 | The Discovery\_Method property is intended to characterize the method and/or tool used to discover the code. |
| **Start\_Address** | cyboxCommon:  HexBinaryObjectPropertyType | 0..1 | The Start\_Address property can be used to reference the start address of the code, if it was discovered inside a binary. |
| **Code\_Segment** | cyboxCommon:  StringObjectPropertyType | 0..1 | The Code\_Segment property encompasses any arbitrary code segment in unencoded (plaintext or binary) format. Code would typically be included here within a CDATA section. |
| **Code\_Segment\_XOR** | CodeObj:  CodeSegmentXORType | 0..1 | The Code\_Segment\_XOR property encompasses any arbitrary code segment. Its contents should contain the actual code segment XORed with the pattern defined in the xorpattern property. This is so that the code contained in the pattern does not trigger IDS, AV, or other signature-based scanners. XOR'd Code would typically be included here within a CDATA section. |
| **Digital\_Signatures** | cyboxCommon:  DigitalSignaturesType | 0..1 | The Digital\_Signatures property captures one or more digital signatures for the code. |
| **Extracted\_Features** | cyboxCommon:  ExtractedFeaturesType | 0..1 | The Extracted\_Features property A description of features extracted from this code segment. |

## CodeTypeType Class

CodeTypeType class specifies class of code, via a union of the CodeTypeEnum enumeration and the atomic xs:string type. Its base type is the CybOX Core BaseObjectPropertyType class, for permitting complex (i.e. regular-expression based) specifications. The UML diagram corresponding to the CodeObjectType class is shown in **Figure 3‑2**.

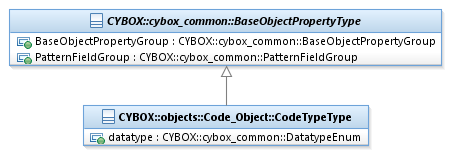


Figure 3‑2. UML diagram of the CodeTypeType class

## CodeSegmentXORType Class

Used to encapsulate a segment of code that has been XORed with a pattern in order to avoid tripping anti-virus detection. The UML diagram corresponding to the CodeSegmentXORType class is shown in **Figure 3‑3**.

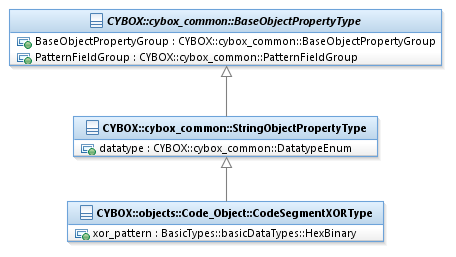


Figure ‑. UML diagram of the CodeSegmentXORType class

The property table of the CodeSegmentXORType class is given in **Table 3‑2**.

Table ‑. Properties of the CodeSegmentXORType class

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Type** | **Multiplicity** | **Description** |
| **xor\_pattern** | xs:hexBinary | 0..1 | The xor\_pattern property contains a 16-hexadecimal-character hex string, which represents the pattern that the Code\_Segment\_XOR field should be XORed with in order to recover the actual code. The default value is 55AA55AA55AA55BB, as specified by IETF RFC 5901. |

## CodePurposeType Class

CodePurposeType class specifies intended purposes of code, via a union of the CodePurposeEnum enumeration and the atomic xs:string type. Its base type is the CybOX Core BaseObjectPropertyType class, for permitting complex (i.e. regular-expression based) specifications. The UML diagram corresponding to the CodePurposeType class is shown in **Figure 3‑4**.

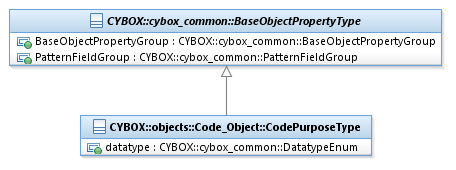


Figure ‑. UML diagram of the CodePurposeType class

## CodeLanguageType Class

CodeLanguageType class specifies languages of code, via a union of the CodeLanguageEnum enumeration and the atomic xs:string type. Its base type is the CybOX Core BaseObjectPropertyType class, for permitting complex (i.e. regular-expression based) specifications. The UML diagram corresponding to the CodeLanguageType class is shown in **Figure 3‑5**.

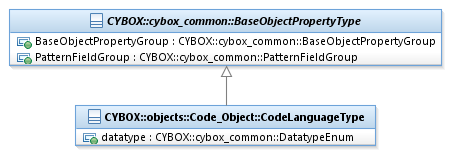


Figure ‑. UML diagram of the CodeLanguageType class

## ProcessorTypeType Class

ProcessorTypeType class specifies relevant processor families, via a union of the ProcessorTypeEnum enumeration and the atomic xs:string type. Its base type is the CybOX Core BaseObjectPropertyType class, for permitting complex (i.e. regular-expression based) specifications. The UML diagram corresponding to the ProcessorTypeType class is shown in **Figure 3‑6**.

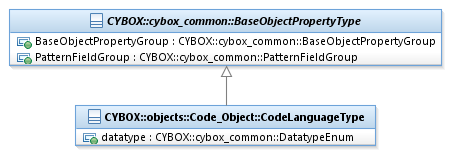


Figure ‑. UML diagram of the ProcessorTypeType class

## TargetedPlatformsType Class

A list of targeted platforms.

The property table of the TargetedPlatformsType class is given in **Table 3‑3**.

Table ‑. Properties of the TargetedPlatformsType class

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Type** | **Multiplicity** | **Description** |
| **Targeted\_Platform** | cyboxCommon:  PlatformSpecificationType | 1..\* | The Targeted\_Platform property specifies a particular platform that this code is targeted for. |

## CodeTypeEnum Enumeration

The literals of the CodeTypeEnum enumeration are given in **Table 3‑4**.

Table ‑. Literals of the CodeTypeEnum enumeration

|  |  |
| --- | --- |
| **Enumeration Literal** | **Description** |
| **Source\_Code** | The code represented is in the form of Source Code. |
| **Byte\_Code** | The code represented is in the form of Byte Code. |
| **Binary\_Code** | The code represented is in the form of binary code. |

## CodePurposeEnum Enumeration

The literals of the CodePurposeEnum enumeration are given in **Table 3‑5**.

Table ‑. Literals of the CodePurposeEnum enumeration

|  |  |
| --- | --- |
| **Enumeration Literal** | **Description** |
| **Application\_Code** | The code represented is intended as application code. |
| **Library\_Code** | The code represented is intended as library code. |
| **Shellcode** | The code represented is intended as shellcode. |
| **Exploit\_Code** | The code represented is intended as exploit code. |
| **Unknown** | The code represented is intended for unknown purposes. |
| **Other** | The code represented is intended for a purpose other than those listed in this enumeration. |

## CodeLanguageEnum Enumeration

The literals of the CodeLanguageEnum enumeration are given in **Table 3‑6**.

Table ‑. Literals of the CodeLanguageEnum enumeration

|  |  |
| --- | --- |
| **Enumeration Literal** | **Description** |
| **C** | Indicates the code is written in the C programming language. |
| **C++** | Indicates the code is written in the C++ programming language. |
| **C#** | Indicates the code is written in the C# programming language. |
| **Java** | Indicates the code is written in the Java programming language. |
| **JSP** | Indicates the code is written in the JSP (Java Server Pages) language. |
| **Javascript** | Indicates the code is written in the Javascript programming language. |
| **ASP.NET** | Indicates the code is written in the ASP.NET programming language. |
| **SQL** | Indicates the code is written in SQL (Standard Query Language). |
| **Python** | Indicates the code is written in the Python programming language. |
| **Perl** | Indicates the code is written in the Perl programming language. |
| **PHP** | Indicates the code is written in the PHP programming language. |
| **SOAP** | Indicates the code is written as a SOAP message. |
| **Ruby** | Indicates the code is written in the Ruby programming language. |
| **Shell** | Indicates the code is written as a Shell script. |
| **PseudoCode** | Indicates the code is written as pseudo code. |
| **.NET** | Indicates the code utilizes the .NET framework. |
| **Assembly** | Indicates the code is written in an assembly language. |
| **XML** | Indicates the code is written in XML (eXtensible Markup Language). |
| **HTML** | Indicates the code is written in HTML (HyperText Markup Language). |
| **Other** | Indicates the code is written in a language not found in this enumeration. |

## ProcessorTypeEnum Enumeration

The literals of the ProcessorTypeEnum enumeration are given in **Table 3‑7**.

Table ‑. Literals of the ProcessorTypeEnum enumeration

|  |  |
| --- | --- |
| **Enumeration Literal** | **Description** |
| **x86-32** | Indicates a x86 32bit processor. |
| **x86-64** | Indicates a x86 64bit processor. |
| **IA-64** | Indicates an IA (Intel Itanium) 64bit processor. |
| **PowerPC** | Indicates a PowerPC processor. |
| **ARM** | Indicates an ARM processor. |
| **Alpha** | Indicates an Alpha processor. |
| **SPARC** | Indicates a SPARC processor. |
| **z/Architecture** | Indicates a z/Architecture (IBM) processor. |
| **eSi-RISC** | Indicates an eSi-RISC processor. |
| **MIPS** | Indicates a MIPS processor. |
| **Motorola 68k** | Indicates a Motorola 68k processor. |
| **Other** | Indicates a processor outside of this enumeration. |

# Conformance

Implementations have discretion over which parts (components, properties, extensions, controlled vocabularies, etc.) of CybOX they implement (e.g., Observable/Object).

[1] Conformant implementations must conform to all normative structural specifications of the UML model or additional normative statements within this document that apply to the portions of CybOX they implement (e.g., implementers of the entire Observable class must conform to all normative structural specifications of the UML model regarding the Observable class or additional normative statements contained in the document that describes the Observable class).

[2] Conformant implementations are free to ignore normative structural specifications of the UML model or additional normative statements within this document that do not apply to the portions of CybOX they implement (e.g., non-implementers of any particular properties of the Observable class are free to ignore all normative structural specifications of the UML model regarding those properties of the Observable class or additional normative statements contained in the document that describes the Observable class).

The conformance section of this document is intentionally broad and attempts to reiterate what already exists in this document.

Acknowledgments

The following individuals have participated in the creation of this specification and are gratefully acknowledged:

Participants:

Dean Thompson, Australia and New Zealand Banking Group (ANZ Bank)

Bret Jordan, Blue Coat Systems, Inc.

Adnan Baykal, Center for Internet Security (CIS)

Liron Schiff, Comilion (mobile) Ltd.

Jane Ginn, Cyber Threat Intelligence Network, Inc. (CTIN)

Richard Struse, DHS Office of Cybersecurity and Communications (CS&C)

Ryusuke Masuoka, Fujitsu Limited

Eric Burger, Georgetown University

Jason Keirstead, IBM

Paul Martini, iboss, Inc.

Jerome Athias, Individual

Sanjiv Kalkar, Individual

Terry MacDonald, Individual

Alex Pinto, Individual

Patrick Maroney, Integrated Networking Technologies, Inc.

Wouter Bolsterlee, Intelworks BV

Joep Gommers, Intelworks BV

Sergey Polzunov, Intelworks BV

Rutger Prins, Intelworks BV

Andrei Sîrghi, Intelworks BV

Jonathan Baker, MITRE Corporation

Sean Barnum, MITRE Corporation

Mark Davidson, MITRE Corporation

Ivan Kirillov, MITRE Corporation

John Wunder, MITRE Corporation

Mike Boyle, National Security Agency

Jessica Fitzgerald-McKay, National Security Agency

Takahiro Kakumaru, NEC Corporation

John-Mark Gurney, New Context Services, Inc.

Christian Hunt, New Context Services, Inc.

Andrew Storms, New Context Services, Inc.

Igor Baikalov, Securonix

Bernd Grobauer, Siemens AG

John Anderson, Soltra

Trey Darley, Soltra

Paul Dion, Soltra

Brandon Hanes, Soltra

Ali Khan, Soltra

The authors would also like to thank the larger CybOX Community for its input and help in reviewing this document.

Revision History

|  |  |  |  |
| --- | --- | --- | --- |
| **Revision** | **Date** | **Editor** | **Changes Made** |
| wd01 | 16 November 2015 | Desiree Beck Trey Darley Ivan Kirillov Rich Piazza | Initial transfer to OASIS template |