## **Guidelines for Using the SWIM Controlled Vocabulary**

Carol Uri (The SEMCON Group, LLC), March 2019

What is a Controlled Vocabulary (CV)?

Why do we need a SWIM CV?

What is the source of the SWIM CV terms and definitions?

Why are there HTML and RDF versions of the SWIM CV?

What is SKOS?

How do we link to terms in the SWIM CV?

How can we connect the SWIM CV with other CVs?

How are SWIM CV terms uniquely identified?

How is the SWIM CV documented?

How is each SWIM CV term documented?

- Rules for creating preferred labels
- Rules for writing good definitions
  - Requirements
  - Recommendations
- > Rules for documenting the source of a definition
- Rules for using SKOS relationships
- > Rules for using SKOS notes

## What is a Controlled Vocabulary (CV)?

A *controlled vocabulary* (CV) is a managed list of terms (concepts) that have been enumerated explicitly. All of the terms in a CV have an unambiguous, non-redundant definition and are connected through Uniform Resource Identifiers (URIs) on the Web.

(Back to top)

#### Why do we need a SWIM CV?

A common, agreed-upon vocabulary is a necessary prerequisite for successful communication – be it written, oral or electronic – among technically diverse and independently managed organizations such as SWIM developers and providers. Each organization tends to create its own vocabulary based on engineering and business practices. The usual approach toward eliminating semantic incompatibilities is establishing a single vocabulary to which all potential stakeholders agree to adhere.

Without a CV, SWIM stakeholders must rely on published documents, especially documents with glossaries, to provide terms and definitions they need. Some of the problems with relying on these sources include:

- The terms and definitions are in human-readable documents only; they are not Web-based.
- Terms are not maintained separately, e.g., in a managed dictionary.
- Relationships among the terms are not apparent.
- People who need definitions for some particular terms may not be aware they have already been defined, so they duplicate the effort by redefining them unnecessarily.

To address these problems, SWIM has followed the approach defined and specified in the <u>SKOS (Simple Knowledge Organization System</u>) specification to produce a CV with a Web-based (Semantic Web) structure that can be used to link or map concepts across the Web – and, if needed, can also be rendered as a human-readable document.

(Back to top)

## What is the source of the SWIM CV terms and definitions?

In 2013, the US Federal Aviation Administration's SWIM program developed a vocabulary of common terms (concepts) used for describing and governing Web-based services. Definitions for these concepts were adopted from industry standards bodies wherever possible. These common terms have been used to help create the first version of the SWIM CV.

It should be noted that many of the concepts in the SWIM CV are also concepts in the Service Description Conceptual Model (SDCM) 2.0, a collaborative effort of the FAA/Single European Sky Air Traffic Management (ATM) Research Programme (SESAR) Joint Undertaking (SJU).

(Back to top)

## Why are there HTML and RDF versions of the SWIM CV?

The <u>SWIM CV Web page</u> is an HTML rendition of the actual <u>SWIM CV</u>, which is an <u>RDF</u>-based document developed in accordance with the SKOS specification. The HTML document, which is made human-readable via Web browser applications like Internet Explorer, Chrome, or Firefox, was generated by applying an <u>XSL</u> script to the SWIM CV RDF document. The RDF document is machine-readable, and other software applications may be developed to utilize the CV in other ways. (Note: the RDF document is written using <u>Turtle</u>, a compact and natural text form of RDF.)

(Back to top)

#### What is SKOS?

The SKOS Primer states, "The Simple Knowledge Organization System (SKOS) is an RDF vocabulary for representing semi-formal knowledge organization systems (KOSs), such as thesauri, taxonomies, classification schemes and subject heading lists. Because SKOS is based on the Resource Description Framework (RDF), these representations are machine-readable and can be exchanged between software applications and published on the World Wide Web. [...] In basic SKOS, conceptual resources (concepts) can be identified with URIs, labeled with lexical strings in one or more natural languages, documented with various types of notes, semantically related to each other in informal hierarchies and association networks, and aggregated into concept schemes." Since CV terms are concepts that are intended to be shared, the SWIM CV is implemented in SKOS.

(Back to top)

#### How do we link to terms in the SWIM CV?

To link to the HTML description of the SWIM CV in your documentation, use the address <a href="http://semantics.aero/pages/swim-vocabulary">http://semantics.aero/pages/swim-vocabulary</a>. To link to the HTML description of a specific CV term, use the address <a href="http://semantics.aero/pages/swim-vocabulary/#term">http://semantics.aero/pages/swim-vocabulary/#term</a>. For example, to link to the term "service", use <a href="http://semantics.aero/pages/swim-vocabulary/#service">http://semantics.aero/pages/swim-vocabulary/#service</a>.

Note: if a term consists of more than one word, such as "service provider", replace spaces between the words in a term with hyphens, like this: <a href="http://semantics.aero/pages/swim-vocabulary#service-provider">http://semantics.aero/pages/swim-vocabulary#service-provider</a>.

(Back to top)

#### How can we connect the SWIM CV to other CVs?

An important reason for implementing the SWIM CV using SKOS is to be able to relate, or map, SWIM CV terms to terms that come from other CVs on the Web. As the <u>SKOS Primer</u> says, "Every SKOS concept is assigned a URI, which makes it possible to unambiguously reference a concept in any SKOS

application... A crucial feature of mapping is the possibility to state that two concepts from different schemes have comparable meanings, and to specify how these meanings compare, even though they come from different contexts and possibly follow different modeling principles." This mapping is accomplished through the use of attributes like **skos:exactMatch**, **skos:broader**, **skos:narrower**, and **skos:related** as described in the <u>rules for using SKOS relationships</u>. For example, the fact that the SWIM CV term "service provider" (URI http://semantics.aero/swim-vocabulary#service-provider) has the same meaning as the World Wide Web Consortium (W3C) Web Services Glossary term "service provider" (URI http://www.w3.org/TR/ws-gloss/#provider) is shown by adding a skos:exactMatch attribute to the SWIM CV:

```
skos:exactMatch <a href="http://www.w3.org/TR/ws-gloss/#provider">http://www.w3.org/TR/ws-gloss/#provider</a>;
```

In addition to mapping to terms in different CVs, there is also the possibility of extending or growing the SWIM CV by adding references to other CVs. As stated in the SKOS Primer, "Linking concepts by means of mappings is not the only way to interlink concept schemes. The use of URIs on the Semantic Web allows resources to be shared and reused in a distributed fashion. As a result it is possible for a SKOS concept to participate in several concept schemes at the same time. For example, a SKOS publisher can choose to locally extend an existing concept scheme by declaring any new concepts that may be needed and simply linking to concepts that have already been defined in the existing scheme." For examples on how to do this, see the SKOS Primer Section 3.2 Reusing and Extending Concept Schemes.

(Back to top)

## How are SWIM CV terms uniquely identified?

In SKOS, a CV term is a concept (**skos:Concept**) and is uniquely identified by a Uniform Resource Identifier (URI), enabling anyone to refer to them unambiguously from any context, and making them a part of the World Wide Web. Each CV term's URI is formed by combining the unique URI of the primary resource (the SWIM CV), "http://semantics.aero/swim-vocabulary", with the fragment identifier of the subordinate resource (the term), e.g., "#service-provider".

#### **Example**

Preferred label: service-oriented architecture

URI: http://semantics.aero/swim-vocabulary#service-oriented-architecture

(Back to top)

#### How is the SWIM CV documented?

The SWIM CV itself is a resource on the Web, and so is each of the terms within the CV. Since it is a Web resource, the CV is documented using selected Dublin Core (dc) standard attributes as follows:

```
<http://semantics.aero/swim-vocabulary>
       rdf:type
                       owl:Ontology;
                      "Mark Kaplun (FAA)",
       dc:creator
                      "Carol Uri (The SEMCON Group, LLC)",
                       "Wen Zhu (NIRA, Inc.)";
       dc:description "The purpose of the SWIM Controlled Vocabulary (CV) is to give
SWIM organizations, support contractors, vendors, and business partners a uniform
understanding of terms employed in the SWIM environment. The CV contains a
comprehensive list of terms with clear and unambiguous definitions. Each term is
globally uniquely identified by a dereferenceable URI so that it can be related
semantically to other terms, vocabularies, or resources.
                     "RDF/SKOS, HTML" ;
       dc:format
       dc:publisher "FAA SWIM";
```

#### (Back to top)

#### How is each SWIM CV term documented?

Each term listed in the SWIM CV is documented using a predefined set of attributes, all but one defined in the SKOS namespace (context). Attributes from the Dublin Core namespace can also be used, though at present there is only one, which is the definition's source ("dc:source").

Below are the attributes that may be used to document a CV term. Each attribute is hyperlinked to the section of the SKOS Primer that explains it further. A CV term always has one unique URI and one preferred label; the other attributes are optional.

CV Term Attribute	SKOS Expression	Meaning
URI (required)	skos:Concept	The term's uniform resource identifier.
Preferred Label (required)	skos:prefLabel	The expression normally used to refer to the term in natural language. <i>More guidance on labeling</i>
Alternative Label	skos:altLabel	A synonym, near-synonym, acronym or abbreviation that is also used to refer to the term.
Definition	skos:definition	A statement of the meaning of the term. <u>More guidance on how</u> to write good definitions
Source of Definition	<u>dc:source</u>	The network location of the document or other resource from which the term's definition is obtained. <u>More guidance on documenting the source of a definition</u>
Editorial Note	skos:editorialNote	Generally, the name of the document or other resource from which the term's definition is obtained.
Concept Scheme	skos:inScheme	The URI of the resource containing the term; here, the SWIM Vocabulary URI (http://semantics.aero/swim-vocabulary).
Exact Match	skos:exactMatch	A term having the same or equivalent meaning. <u>More guidance</u> on using relationships
Broader	skos:broader	A related term which is broader or more general in meaning.
Narrower	skos:narrower	A related term which is narrower or more specific in meaning.
Related	skos:related	A related term which is associated in some non-hierarchical way.
Change Note	skos:changeNote	Used for administrative purposes. <u>More guidance on using SKOS</u> <u>notes</u>
History Note	skos:historyNote	Describes changes to a term's meaning.
Scope Note	skos:scopeNote	Describes limitations on a term's use.
Note	skos:note	Used for general documentation.

An example of the term "service provider" as it would appear on the SWIM CV Website is shown below.

## service provider

URI:	JRI: http://semantics.aero/swim-vocabulary#service-provider	
Preferred Label:	service provider	
Alternative Label:	provider entity	
Definition:	An organization that offers the use of capabilities by means of a service.	
Source of Definition:	http://docs.oasis-open.org/soa-rm/v1.0/soa-rm.pdf	
Editorial Note:	Name of source: OASIS Reference Model for SOA 1.0, 12 October 2006	
<b>Concept Scheme:</b>	http://semantics.aero/swim-vocabulary	
Exact Match:	http://www.w3.org/TR/ws-gloss/#provider	
Broader:	organization	
Related:	service consumer	
Related:	<u>service</u>	

Here is the same term rendered in SKOS/RDF using Turtle, a compact form of RDF:

#### (Back to top)

## Rules for creating preferred labels

Preferred labels (**skos:prefLabel**) are used to hold the expression normally used to refer to the CV term. There can be only one preferred label per language. (Note: the SWIM CV only contains English terms at this time.) Rules for creating preferred labels are:

**1.** Capitalization – Begin all words in a multi-word preferred label with *lower-case letters* unless they are proper nouns (proper noun: a specific individual, place, etc., that is not normally used with an article, and that is normally capitalized) or proper adjectives (proper adjective: an adjective formed from a proper noun, e.g., "Italian").

#### **Example**

*Correct*: service consumer, real world effect, Web service ("Web" is normally capitalized) *Incorrect*: Service Consumer, Real World Effect, Web Service

**2. Terms with more than one definition** – If a term has more than one English definition, i.e., if it means different things depending on the context in which it is used, create *separate terms* for each definition by including a context as part of the term's preferred label.

## Example

Correct:

- (1) service provider An organization that offers the use of capabilities by means of a service.
- (2) JMS provider A messaging system that implements the JMS API in addition to the other administrative and control functionality required of a full-featured messaging product

provider – An organization that offers the use of capabilities by means of a service; a messaging system that implements the JMS API in addition to the other administrative and control functionality required of a full-featured messaging product.

**3. Acronyms** – Although a term's acronym or other abbreviation may often be used in the SWIM environment as a substitute for the fully spelled out term, like "SOA" or "MOM" (or "SWIM" for that matter), best practice is to use the full term as the preferred label and the acronym as its alternative label. Only acronyms like "radar" (radio detecting and ranging) or "laser" (light amplification by stimulated emission of radiation) that have entered our English language do not need to be spelled out.

## **Example**

Preferred label: service-oriented architecture Alternative label: SOA

**4.** When a CV term's preferred label contains the acronym of another CV term used as a modifier, like "SOA governance", the acronym does not need to be spelled out in the preferred label. The term's documentation should include a relationship to the other term, however.

#### **Example**

Preferred label: SOA governance Related: service-oriented architecture

(Back to attributes)

(Back to top)

## Rules for writing good definitions

**Definition**: A word or phrase expressing the essential nature of a person or thing or class of persons or things; an answer to the question "what is x?" or "what is an x?"; a statement of the meaning of a word or word group. [Webster's Third New International Dictionary of the English Language Unabridged, 1986]

The purpose of a definition is to define a concept with words or phrases that describe, explain, or make definite and clear its meaning. *Precise* and *unambiguous* definitions are one of the most critical aspects of ensuring interoperability. When two or more parties use a term, it is essential that all be in explicit agreement on the meaning of that term.

<u>ISO/IEC 11179-4</u> provides a guide for writing good definitions. There are mandatory <u>requirements</u> with which all definitions must comply, and there are <u>recommendations</u> that should be followed when writing a definition. Note the difference between rules and guidelines: compliance with the rules can be objectively tested, whereas compliance with the guidelines can only be evaluated subjectively. Many of the rules and guidelines cited below are abstracted from this document.

Although ISO/IEC 11179-4 requirements and recommendations pertain to data concepts and other administered items in a metadata registry, they can also be applied when writing definitions for concepts in general.

## Requirements

A definition *shall*:

- 1. Be stated in the singular.
- 2. State what the concept is, not only what it is not (i.e., never exclusively in the negative).
- 3. Be stated as a descriptive phrase or sentence(s).
- 4. Contain only commonly used abbreviations.
- 5. Be expressed without embedding definitions of other underlying concepts.

Descriptions and examples of each requirement are provided below. (Note: the terms used in these examples are not SWIM Vocabulary terms; they are only for illustration.)

#### 1. Be stated in the singular.

The concept expressed by the definition must be stated in the singular. (An exception is made if the concept itself is plural.)

Example: "article number"

**Good:** A reference number that identifies an article. **Poor:** A reference number that identifies articles.

Reason: The poor definition uses the plural word "articles," which is ambiguous since it could imply that an "article number" refers to more than one article.

2. State what the concept is, not only what it is not (i.e., never exclusively in the negative).

A definition cannot be constructed exclusively by saying what the concept is not.

Example: "freight cost"

**Good:** Cost incurred by a shipper in moving goods from one place to another.

**Poor:** Cost not related to packing, documentation, loading, unloading, and insurance.

Reason: The poor definition does not specify what is included in the meaning of the concept.

## 3. Be stated as a descriptive phrase or sentence(s).

A phrase or sentence is necessary to describe the essential characteristics of the concept. Simply restating the concept as a synonym, or restating it with the same words, is not sufficient. If more than one descriptive phrase is needed, use complete grammatically correct sentences.

Example: "weather forecast"

**Good**: An estimation or calculation of future weather conditions.

**Poor**: A weather prediction.

Reason: "Weather prediction" is just a near-synonym for the name of the concept, which is not adequate for a definition.

#### 4. Contain only commonly understood abbreviations.

Understanding the meaning of an abbreviation or acronym is usually confined to a certain environment. In other environments, the same abbreviation can cause misinterpretation or confusion. Exceptions may be made for common abbreviations such as "i.e." and "e.g." or if an abbreviation is more readily understood than the full form and has been adopted as a term in its own right, such as "radar" (radio detecting and ranging). When an acronym is first used in a definition, it should be expanded.

Example: "elevation"

**Good:** The vertical distance of a point or a level on, above, or below the surface of the earth, measured from the earth's mean sea level (MSL) datum.

**Poor:** The vertical distance from MSL to a specific point.

Reason: The poor definition is unclear because the acronym MSL is not commonly understood and some users may need to determine what it represents.

## 5. Be expressed without embedding definitions of underlying concepts.

The definition of a second concept should not appear in the definition proper of the primary concept. Definitions of terms should be provided separately.

Example: "aircraft damage code"

**Good:** A code that designates the level of damage sustained by an aircraft as a result of an accident.

**Poor:** A code that designates the level of damage sustained by the aircraft as a result of an accident. An aircraft accident is an occurrence associated with the operation of an aircraft that takes place between the time any person boards the aircraft with the intention of flight and the time all such persons have disembarked, and in which any person suffers death or serious injury, or in which the aircraft receives substantial damage.

Reason: The poor definition contains an extraneous definition embedded in it, which is the definition of "aircraft accident".

(Back to attributes)

(Back to top)

#### Recommendations

A definition *should*:

- 1. State the essential meaning of the concept.
- 2. Be precise and unambiguous.
- 3. Be concise.
- 4. Be able to stand alone.
- 5. Be expressed without embedding rationale, functional usage, domain information, or procedural information.
- 6. Avoid circular reasoning.
- 7. Use the same terminology and consistent logical structure for related definitions.

Descriptions and examples of each recommendation are provided below. (Note: the terms used in these examples are not SWIM Vocabulary terms; they are only for illustration.)

## 1. State the essential meaning of the concept.

Include all primary aspects of the concept, but avoid non-essential characteristics.

**Example**: "invoice amount"

**Good:** The total sum charged on an invoice.

**Poor:** The total sum of all chargeable items mentioned on an invoice, taking into account deductions on one hand, such as allowances and discounts, and additions on the other hand, such as charges for insurance, transport, handling, etc.

Reason: The poor definition includes extraneous material.

## 2. Be precise and unambiguous.

The exact meaning and interpretation should be apparent from the definition. A definition should be clear enough to allow only one possible interpretation.

**Example:** "shipment receipt date"

**Good:** The date on which a shipment is received by the receiving party.

**Poor:** The date on which a specific shipment is delivered.

Reason: The poor definition does not specify what determines a "delivery." "Delivery" could be understood as either the act of unloading a product at the intended destination or the point

at which the intended customer actually obtains the product. It is possible that the intended customer never receives the product that has been unloaded at his site or the customer may receive the product days after it was unloaded at the site.

#### 3. Be concise.

The definition should be brief and comprehensive. Extraneous qualifying phrases such as "terms to be described" or "for the purposes of" are to be avoided. The definition should not begin with an expression such as "term used to describe" or "term denoting," nor should it take the form "is...," "means...," "one of...".

**Example**: "NCP number"

**Good:** A unique identifier assigned to a National Airspace System (NAS) Change Proposal (NCP) case file by the NAS Configuration Control Board.

**Poor:** The NCP number is a unique identifier assigned to a National Airspace System (NAS) Change Proposal (NCP) case file by the NAS Configuration Control Board for the purpose of NAS CCB administrative procedures or for use in retrieving case files from the NAS CCB's automated configuration management system, WebCM.

Reason: In the poor definition, the name of the concept is repeated ("The NCP number is..."), and the phrases after "...Control Board" are extraneous qualifying phrases.

#### 4. Be able to stand alone.

The meaning of the concept should be apparent from the definition. Additional explanations or references should not be necessary to understand the meaning of the definition.

Example: "accident location city"

**Good:** The name of the city nearest to the accident site.

**Poor:** See "event site" in FAA Order 8020.11C.

Reason: The poor definition does not stand alone, but requires the aid of a second definition (event site) to understand the meaning of the first.

# 5. Be expressed without embedding rationale, functional usage, domain information, or procedural information.

Reasons as to why the definition is expressed a certain way should not be included in the definition. Functional usage (e.g., "this term should not be used for...") or procedural aspects (e.g., "this term is used in conjunction with ...") are more properly handled as comments or related references.

Example: "midair collision indicator"

**Good:** A code that indicates whether or not an accident involved a midair collision between two aircraft.

**Poor:** A code that indicates whether or not an accident involved a midair collision between two aircraft. This code is used to count collisions in the air, not on the ground and not with objects (towers).

Reason: Remarks about functional usage (i.e., "this code is used to count...") should be omitted from the definition. If this information is needed, it should be entered as a comment.

#### 6. Avoid circular reasoning.

Two concepts should not be defined in terms of each other. A concept should not use the definition of another concept as its definition.

**Example**: two concepts, "employee" and "employee identification number"

**Poor:** (1) employee – A person who has been assigned an employee identification number.

(2) employee identification number – A number assigned to an employee.

Reason: Each definition refers to the other for its meaning. The meaning is not given in either definition.

7. Use the same terminology and consistent logical structure for related definitions.

Use common terminology and syntax (i.e., consistent logical structure) for similar or associated definitions to facilitate understanding.

Example: two concepts, "goods dispatch date" and "goods receipt date"

**Good:** (1) goods dispatch date – The date on which goods were dispatched by a given party.

(2) goods receipt date – The date on which goods were received by a given party.

**Poor:** (1) goods dispatch date – The date on which goods were dispatched by a given party.

(2) goods receipt date – The date on which the customer received the merchandise.

Reason: Users may wonder whether some difference is implied by the use of synonymous terms and variable syntax.

(Back to attributes)

(Back to top)

## Rules for documenting the source of a definition

- **1.** The **dc:source** attribute is used to hold the network location (URL) of the resource from which the definition was taken. If there is no network location, **omit** the dc:source attribute.
- **2.** The **skos:editorialNote** attribute is used to hold the bibliography of the resource from which the definition was copied or derived. Because this attribute may be used for other purposes as needed, the convention is to begin with the phrase "**Name of source:**" and follow that with the bibliography (resource title, version, publisher, date, etc.) For example:

```
skos:prefLabel "service provider" skos:definition "An organization that offers the use of capabilities by means of a service." dc:source "<a href="http://docs.oasis-open.org/soa-rm/v1.0/soa-rm.pdf">http://docs.oasis-open.org/soa-rm/v1.0/soa-rm.pdf</a>" skos:editorialNote "Name of source: OASIS Reference Model for SOA 1.0, 12 October 2006"
```

**3.** If the definition was not copied or derived from a resource, i.e., if it was developed in-house, the convention is to begin with the phrase "**Definition developed by**" and follow that with the name of the developing organization. For example:

```
skos:prefLabel "message producer" skos:definition "An application or process that creates and sends messages." skos:editorialNote "Definition developed by FAA SWIM"
```

(Back to attributes)

(Back to top)

#### **Rules for using SKOS relationships**

- **1.** The attributes **skos:exactMatch**, **skos:broader**, **skos:narrower**, and **skos:related** are used in the SWIM CV to hold the URIs of any terms which are related to the CV term being documented. If there are no related terms, these attributes are **omitted**. There is no theoretical limit to the number of related terms that can be included.
- **2. skos:exactMatch** A term that is an "exact match" to a SWIM CV term being documented will always come from a **different vocabulary**. This is because if such a term were in the SWIM CV, it would just be an alternative label (skos:altLabel) for the preferred SWIM CV term.
- **3.** skos:broader/skos:narrower If a term X has a related term Y which is broader or more general, then X is implicitly narrower than Y; that is, it is not necessary also to define term Y as having a narrower term X. For example, if "service provider" (skos:prefLabel "service provider") is related to the broader term "organization" (skos:broader "organization"), there is no need to document the term "organization" as

having a related narrower term "service provider". For more information, see the <u>SKOS Primer Section</u> 2.3 Semantic Relationships.

(Back to attributes)

(Back to top)

## **Rules for using SKOS notes**

Other attributes that may be used to document SWIM CV terms are shown below. If there is no need to include such information, these attributes are **omitted**. There is no theoretical limit to the number of notes that can be included.

- **1.** skos:scopeNote Supplies information on the intended meaning or limitations of a term.
- **2. skos:historyNote** Describes significant changes to a term's meaning or form.
- **3.** skos:changeNote Used for maintenance and administration, e.g., approval status and date.
- **4. skos:note** Used for any other general documentation purposes.

For more information, see the **SKOS** Primer Section 2.4 Documentary Notes.

(Back to attributes)

(Back to top)