|  |  |  |
| --- | --- | --- |
| SNOMED CT URI Specification | | |
|  | | |
| Date | 2013???? | |
| Version | | 0.07 |

Amendment History

|  |  |  |  |
| --- | --- | --- | --- |
| **Version** | **Date** | **Editor** | **Comments** |
| 0.01 | 2012 08 07 | Michael Lawley | Initial document |
| 0.02 | 2012 09 04 | Michael Lawley | Addition of detail surrounding Extensions and Releases |
| 0.03 | 2012 10 02 | Michael Lawley | Revised to address feedback from Technical Committee |
| 0.04 | 2013 01 07 | Michael Lawley | Further revisions addressing feedback from Technical Committee and addition of role grouping concept |
| 0.05 | 2013 03 04 | Michael Lawley | Added impact statement and additional refinements based on feedback from Technical Committee. |
| 0.06 | 2013 04 30 | Michael Lawley | Revised following QA review feedback. This included a statement of who should read this document, clarification that it defines a standard and is not a guideline, and a clarification that the "Additional URIs" of section 5 can be more clearly characterised as "URIs for Properties". |
| 0.07 | 2013 ?? ?? | Michael Lawley | Revised following CoP feedback. |
|  |  |  |  |

Review Timetable

This section is populated once final approval is received to indicate when the content will be formally reviewed – date is agreed by Management Board

|  |  |  |
| --- | --- | --- |
| **Review date** | **Responsible owner** | **Comments** |
| YYYYMMDD | Person/group responsible | Summary of action |
|  |  | (remove or add rows if necessary) |

© International Health Terminology Standards Development Organisation 2013. All rights reserved.

SNOMED CT® was originally created by the College of American Pathologists.

This document forms part of the International Release of SNOMED CT® distributed by the International Health Terminology Standards Development Organisation (IHTSDO), and is subject to the IHTSDO’s SNOMED CT® Affiliate Licence. Details of the SNOMED CT® Affiliate Licence may be found at *http://www.ihtsdo.org/our-standards/licensing/*.

No part of this document may be reproduced or transmitted in any form or by any means, or stored in any kind of retrieval system, except by an Affiliate of the IHTSDO in accordance with the SNOMED CT® Affiliate Licence. Any modification of this document (including without limitation the removal or modification of this notice) is prohibited without the express written permission of the IHTSDO.

Any copy of this document that is not obtained directly from the IHTSDO [or a Member of the IHTSDO] is not controlled by the IHTSDO, and may have been modified and may be out of date. Any recipient of this document who has received it by other means is encouraged to obtain a copy directly from the IHTSDO [or a Member of the IHTSDO. Details of the Members of the IHTSDO may be found at *http://www.ihtsdo.org/members/*].

Table of Contents

1 Introduction 5

1.1 Purpose of document 5

1.2 Who Should Use this Specification? 5

1.3 Scope of document 5

1.4 Motivating factors 6

1.5 Use cases for URIs 7

1.6 A note on Releases and Editions 8

1.7 Statement of Impact 9

2 URIs for Components 10

3 URIs for Editions 11

4 URIs for Modules 12

5 URIs for Properties 13

# Introduction

## Purpose of document

This document describes a specification for a standard format of URIs for identifying various SNOMED CT artefacts including components and RF2-based releases.

As a specific sub-case this includes URIs for formally identifying SNOMED CT and any of its extensions, as well as a specific *Version* thereof.

## Who Should Use this Specification?

The intended audience for this document includes both technical professionals who are involved in the development or implementation of terminology systems or healthcare information systems that use SNOMED CT, as well as academics, researchers, and others who are using SNOMED CT in the context of OWL and other Semantic Web technologies. This specification should be used in cases where it is required to uniquely identify SNOMED CT concepts and other components in contexts where URIs are expected, or where the interpretation of a code as an SCTID may be ambiguous. It should also be used when an unambiguous interoperable (machine-readable) identifier for an Edition (or a Version thereof) is required.

For guidance on using the SNOMED CT URI Specification, including resolvability of the URIs, please consult the SNOMED CT URI Guide.

## Scope of document

This document provides a specification for the format and usage of SNOMED CT URIs. This specification relies on the semantics of SNOMED CT modules as defined in the Release Format 2 specification. Please see Section *5.4 Release Format 2 – Core Component Guide*[[1]](#footnote-1) in the separate document “SNOMED CT® Technical Implementation Guide” for additional information on this subject.

The additional document "SNOMED CT URI Guide" provides information on how these identifying URIs may be used in various scenarios including a discussion of the use of the http: URI scheme and approaches to dereferencing the URIs.

## Motivating factors

The existing SCTID specification allows for the identification of a component across time (i.e., the rows in a table that represent the state of that component at a series of points in time). However, this is but one, low level, view of a component. There are other views of a component that are useful to be able to identify. These include, for example, a Concept including its Descriptions and Relationships in a given combination of SNOMED CT International and its Extensions, at a given point in time. Furthermore certain things, such as an Extension with all its dependent modules, are not themselves components, but also need a consistent identification mechanism. This not only includes the individual 6-monthly releases of the International version of SNOMED CT, but also specific national versions such as the Australian release, or the Swedish translation.

A number of groups have emphasised the need to come up with an approach that addresses the broad needs of implementers and offers the opportunities for use of a ubiquitous range of services using the URI as a common factor in the interfaces. This document describes a URI space that is intended to meet these requirements (and to evolve to meet others as they emerge) to avoid the proliferation of alternative conflicting schemes.

The URI space defined in the document uses the syntax defined in *IETF RFC6570 URI Templates*[[2]](#footnote-2). In addition, principals of good URI design were drawn from the W3C document *Cool URIs for the Semantic Web*[[3]](#footnote-3), and *Designing URI Sets for the UK Public Sector*[[4]](#footnote-4).

It should be further noted that, consistent with the advice of Tim Berners-Lee[[5]](#footnote-5), the http scheme is used for these URIs. Furthermore, to be consistent with the W3C's TAG resolution of *ISSUE-14[[6]](#footnote-6)*, since the URIs defined in this document identify *real-world objects* and not *information resources*, resolving these URIs should **not** result in an HTTP response code of 200 ("OK") but rather, if anything at all, result in an HTTP response code of 303 ("See Other") to redirect to another URI that identifies a representation of the identified component. The intuition here is that it is not possible to return a real-world object (e.g., "The Eiffel Tower"), but only a representation of it (a picture, a geo-location, a Wikipedia page, etc.). In the same manner, it is only possible to return a representation of the identified SNOMED CT component, and not the component itself. Further discussion around this issue can be found in Section 4.4 *Choosing between 202 and Hash[[7]](#footnote-7)* of the aforementioned W3C document *Cool URIs for the Semantic Web*.

## Use cases for URIs

The following use cases have guided the specification detailed in this document:

1. The OWL representation of the stated form of SNOMED CT requires URIs to identify Concepts and Object Properties (Attributes). It has historically used its own de facto URI space for this purpose, and has not directly addressed the issue of a URI to identify the ontology itself.
2. The CTS2 specification requires all resources to be identified using URIs. It too has a proposed approach with a narrower scope than we have here.
3. Within the HL7 community there is a need for a consistent mechanism to identify “versions” of SNOMED CT. An appropriate URI space could simply address this need in an extensible fashion.

While a register of canonical names for each Edition could be compiled and maintained, the module system developed for Release Format 2 already provides the required machinery to support unique naming of Editions and, in conjunction with a timestamp, specific versions of an Edition.

Section *5.4.1.4. Identification of Source Module[[8]](#footnote-8)* of the Technical Implementation Guide says the following:

A *moduleId* field, assigned to each component, helps identify the origin of content and dependencies in a release. This enables release centers to compose a unified release from a number of different modules, yet still identify the origin of content within the release. For example, module ids may be used to differentiate *SNOMED CT* International content, Australian Medicines terminology and Pathology content within the Australian national release.

The module dependency reference set is used to track dependencies between (versioned) modules. Thus, by tracing the set of module dependencies from a specified (versioned) moduleId, one is able to identify all the content relevant to that (versioned) moduleId. Hence, a (versioned) moduleId can be used to uniquely identify a (versioned) Edition.[[9]](#footnote-9)

## A note on Releases and Editions

In this document we use the terms *Release* and *Edition* as follows:

**Release**

This is used to mean a concrete set of files that is published by a release centre (including the IHTSDO). This may include any combination of RF2 files, be they full, snapshot or delta, as well as documentation, cross-map files, alternate identifiers, and so forth. In contrast,

**Edition**

This is used to refer to the content of an Extension's modules and all the modules they depend on. That is, the logical SNOMED CT content that is conceptually managed within the Module and EffectiveTime versioning scheme of RF2. In particular, this includes content that pertains specifically to the meaning of Concepts and the contents of Reference Sets. In this latter sense, the term *Edition* is considered, in the context of this document and the associated Guide, to also include the International version of SNOMED CT and not just its derivatives.

In some cases a Release comprises the union of two (or more) parts. For example, SNOMED CT with the addition of medication terminology. In the case that these parts are truly distinct, then distinct URIs can be used to identify them individually. In the case that they are not distinct (that is, there is a dependency with respect to their content), or one part is intended to only be used in conjunction with the other, then this logical dependency should be explicitly managed. The Module (Version) Dependency Reference Set[[10]](#footnote-10) (see Section 1.4) is an appropriate mechanism for doing this and the SNOMED CT URI Guide contains additional discussion of this topic.

## Statement of Impact

This Specification builds on a number of other elements of the IHTSDO SNOMED CT ecosystem. In particular its semantics are dependent on those of RF2 and the module and versioning mechanism.

This Specification defines a standard set of identifiers in the form of URIs. In order to maintain the integrity of the associated URI space, it is highly desirable for the IHTSDO to maintain ownership of the **snomed.info** DNS domain. While not a requirement of this specification, it would be useful if the URIs defined by this Specification, with respect to SNOMED CT Core, were resolvable.

The Perl script that generates OWL (and other) representations of SNOMED CT that name things with URIs should use URIs conforming to this Specification.

# URIs for Components

Components within SNOMED CT are identified using a numeric identifier, the SCTID[[11]](#footnote-11). URIs corresponding to SCTIDs take the following form:

http://snomed.info/id/{sctid}

In addition, entries in a Reference Set are identified by UUIDs. In this case the corresponding URI is:

http://snomed.info/id/{uuid}

# URIs for Editions

A SNOMED CT Edition consists of the complete set of members of one or more Modules.[[12]](#footnote-12) As the Module (Version) Dependency Reference Set[[13]](#footnote-13) (MVDRS) tracks the explicit dependencies between a version of a Module and all the versioned Modules it depends on, a Module Id[[14]](#footnote-14) is a natural identifier for an Edition. When combined with a Timestamp corresponding to a sourceEffectiveTime appearing in the MVDRS, this unambiguously identifies a version of an Edition.

The URIs that identify unversioned and versioned Editions take the following respective forms:

http://snomed.info/sct/{sctid}

http://snomed.info/sct/{sctid}/version/{timestamp}

Note, while it would be possible to extend this pattern to support multiple root Modules, each with their own sourceEffectiveTime, this introduces non-trivial complexities. For example, the Modules they each depend upon may themselves overlap but have different versions (targetEffectiveTime) in which case the implied content would be inconsistent. The SNOMED CT URI Guide contains additional discussion and guidance on this topic.

# URIs for Modules

The previous section defined URIs for the transitive contents of a Module (with respect to module version dependencies). This section defines URIs identifying the content of a specified Module only. As before, these URIs come in two forms:

http://snomed.info/module/{sctid}

http://snomed.info/module/{sctid}/time/{timestamp}

Note, in this case the timestamp is merely referencing a point in time and the contents of a single Module is just a fragment of one or more Editions, hence the timestamp is not considered to be necessarily identifying a version.

# URIs for Properties

There are additional aspects of SNOMED CT that do not have SCTIDs but still need a URI for use cases such as an OWL representation of SNOMED CT (to identify certain annotations) and for parts of CTS2 such as characteristicTypeId. To address these needs we define a general set of URIs identifying the RF2-based properties of Components.

The URI space for these properties follows the pattern:

http://snomed.info/field/{tableName}.{fieldName}

Valid table names are as described for Data Files with respect to the *ContentType element[[15]](#footnote-15)* in the File Naming Conventions for RF2. Note, these URIs identify the property itself, not the value or values that may be associated with the property.

1. <http://www.snomed.org/tig?t=trg2main_title> Section number relative to July 2012 version. [↑](#footnote-ref-1)
2. <http://tools.ietf.org/html/rfc6570> [↑](#footnote-ref-2)
3. Specifically the section *URIs for Real-World Objects* <http://www.w3.org/TR/cooluris/#semweb> [↑](#footnote-ref-3)
4. <http://www.cabinetoffice.gov.uk/sites/default/files/resources/designing-URI-sets-uk-public-sector.pdf> [↑](#footnote-ref-4)
5. Linked Data <http://www.w3.org/DesignIssues/LinkedData.html> [↑](#footnote-ref-5)
6. ISSUE-14 <http://www.w3.org/2001/tag/group/track/issues/14> [↑](#footnote-ref-6)
7. Choosing between 303 and Hash [http://www.w3.org/TR/cooluris - choosing](http://www.w3.org/TR/cooluris#choosing) [↑](#footnote-ref-7)
8. <http://www.snomed.org/tig?t=trg2main_gen_idsource> Section number relative July 2012 version. [↑](#footnote-ref-8)
9. In the case where a release centre has not organized what they consider to be an Edition to correspond to the transitive contents of a single moduleId, a single additional moduleId can be created that depends on the modules that comprise the Edition and then be subsequently used to identify that Edition. Note that it is non-conformant to release only part of a module. [↑](#footnote-ref-9)
10. 900000000000534007 | Module dependency reference set | [↑](#footnote-ref-10)
11. <http://www.snomed.org/tig?t=trg_app_sctid> [↑](#footnote-ref-11)
12. While there may be additional files associated with a release, it is only the Module content which affects the computable meaning of a Concept (i.e., the inferable relationships and subsumption between post coordinated expressions). [↑](#footnote-ref-12)
13. <http://www.snomed.org/tig?t=trg2rfs_spec_module_depend> [↑](#footnote-ref-13)
14. This is the identifier of the Module Concept, as would be used in the Module Dependency Reference Set. It is not acceptable to use the identifier of a Description associated with a Module. [↑](#footnote-ref-14)
15. <http://www.snomed.org/tig?t=fng_contentType> [↑](#footnote-ref-15)