



# SNOMED CT Identifiers Update Guide

Date 24<sup>th</sup> May 2011  
Version 1.0



## Amendment History

Version	Date	Editor	Comments
0.1	2010 11 12	John Gutai	Document created
0.2	2010 12 17	John Gutai	Updated with committee feedback
0.3	2010 12 21	John Gutai	Updated with committee feedback, and separated into separate Specification and Update Guide.
0.4	2011 01 10	John Gutai	Updated following comments from Committee review
0.5	2011 04 19	John Gutai	Updated following community of practice review
1.0	2011 05 24	John Gutai	Approved by Management Board

© International Health Terminology Standards Development Organisation 2010. 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from the IHTSDO [or a Member of the IHTSDO] is not controlled by the IHTSDO, and may have been modified and may be out of date.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.....	4
1.1 Purpose of document.....	4
1.2 Scope of document.....	4
2 The Current Situation .....	4
3 Description of the Problem .....	5
4 Description of the Proposed Change.....	5
5 Impact of the Change .....	7
5.1 Benefits of the Change .....	7
5.2 Cost of the Change.....	7
6 Appendix A - The Current Format of the SCT Identifier.....	9
6.1 SctId Data Type .....	9
6.2 SctId Representation .....	9
6.3 SctIds and Extensions .....	10
6.4 SctId Constraints .....	11
6.5 Check-digit.....	11
6.6 Partition-identifier.....	12
6.7 Extension namespaces.....	13
6.8 Item identifier digits.....	13
6.9 Example SctIds.....	14



# 1 Introduction

## 1.1 Purpose of document

The purpose of this document is to describe proposed changes to the meaning of the SctId field, and to describe the benefits of those changes.

## 1.2 Scope of document

This document provides update guidance on the change to the usage of SNOMED CT identifiers. Please see the separate document “SNOMED CT Identifiers” for a specification of the SctId format and guidance on its usage.

# 2 The Current Situation

SNOMED CT Identifiers are not meaningless numbers, but comprise the following embedded information:

- The Namespace that the component is currently maintained in
- An item identifier (or ComponentID) - a unique identifier within that namespace
- A two digit partition identifier (or PartitionID), with the second digit representing the type of the component (concept, description, etc.) and a first digit representing whether the component is maintained in the International Release or an extension.
- A check digit

A full description of the format and current meaning of SNOMED CT identifiers can be found in the Appendix to this document.



Namespaces are allocated to Members and Affiliates by IHTSDO, and once an organisation owns a Namespace, they are able to create unique SNOMED CT identifiers within that Namespace. At present, when it is necessary to move a component from one Namespace to another, the component must be retired from the first Namespace and a new component with a new SNOMED CT identifier must be created in the new Namespace. Supporting data is also created describing where the relationship has moved to and where it has been moved from. This migration process is fully described in the IHTSDO Technical Reference Guide (TRG).

## 3 Description of the Problem

Migration of a component from a National extension to the International Release causes significant disruption in the authoring environment and can cause even greater disruption in an implementation environment, where SNOMED CT identifiers would potentially change from one release to the next, disrupting system operation and interoperability.

## 4 Description of the Proposed Change

The proposed approach to rectifying this issue is to change the meaning of the Namespace portion of the SCT identifier from the Namespace that the component is currently maintained in to the Namespace that the component was originally created in.

The moduleId field, introduced in RF2 and held against each SNOMED CT component, would then record the organisation currently responsible for authoring the component, and would therefore not necessarily have the same Namespace as the component itself.



The Partition identifier part of the SctId would also change meaning. The second digit would still represent the type of the component (concept, description, etc.), but the first digit would now simply identify the format of the SctId, either being a short format SctId (not including a namespace identifier) or a long format SctId (including a namespace identifier). Only IHTSDO can issue a short format SctId (without a namespace), whereas an extension owner can issue a long format SctId (using one of their namespace identifiers).

Following the change, migration of components between extensions would be possible without a change to their SctIds, according to the following rules:

- A component could move from any extension to the International release without a change to its SctId.
- A component could also move from an extension to a parent extension without a change to its SctId.
- A component could move from the International release back to its originating extension without a change to its SctId.
- In all other cases, the existing rules for moving components between extensions should be used and a change of SctId would be necessary.

In order to make explicit which extensions are parents of which other extensions, concepts under the “Namespace Concept” may now form a sub-hierarchy of namespaces. All namespaces at the top level of this hierarchy are considered to have as their parent the International release.

Guidance has also been developed for producers and consumers of SctIds, to help avoid conflicts of ownership and to facilitate identification of owning organisations. These are fully described in the “SNOMED CT Identifiers” document.

This amendment would require changes to the SNOMED CT release process, tooling, and an update to the SNOMED CT Technical Reference Guide. These changes will be planned and implemented during a six month implementation period following approval of this change by the Management Board.



## 5 Impact of the Change

### 5.1 Benefits of the Change

The key benefits of making this change are:

- Large scale retirement and replacement of SctIds place an increased maintenance burden on implementers with no perceivable benefit. This change significantly reduces that burden.
- The change maintains the distinction of the namespace and module identifiers - the former for the creators of content and the latter for the maintainers.
- The change eases the burdens of content providers in the chain of submissions to National extension and the IHTSDO in detecting their content in public releases. It enables them to set policies on how to detect and manage content migration.
- Long term contributions will come from existing extensions. This change will reduce impact on both the National Release Centre extension managers and the source providers.
- The change removes the disincentive to migrate content to the International release or to a parent extension.
- It will enable more frequent incremental release of content due to decreased migration burden.

### 5.2 Cost of the Change

Total cost to IHTSDO is expected to be US\$20K + 10 days officers' time. This estimate covers any required changes to tooling and updates to the documentation.

In addition, there are likely to be costs to those who have implemented SNOMED CT and to suppliers. These costs should be more fully evaluated by Members and Affiliates during the consultation process, but may include:



- Any costs associated with relaxing existing validation rules for SctIds when importing extensions from a releasing authority.
- Costs associated with implementation of any additional validation rules, in line with the guidelines detailed in the “SNOMED CT Identifiers” document.
- Tooling costs to Members, associated with any required changes to the automation of the process to manage component promotion to the International release.



## 6 Appendix A - The Current Format of the SCT Identifier

Components within SNOMED Clinical Terms are identified using numeric identifiers. In the descriptions of the individual tables these identifiers are noted as having the data type SctId. This section describes the characteristics of all fields with the data type SctId.

These identifiers have a common set of characteristics and obey a set of rules, which enable each identifier to refer unambiguously to a single component.

### 6.1 SctId Data Type

The SctId data type is a 64-bit integer, which is subject to the following constraints:

- Only positive integer values are permitted.
- Leading zeros are not permitted.
- The minimum permitted value is 100,000 (6 digits)
- The maximum permitted value is 999,999,999,999,999 (18-digits).
- As a result of rules for the partition-identifier and check-digit, many integers within this range are not valid SctIds.

### 6.2 SctId Representation

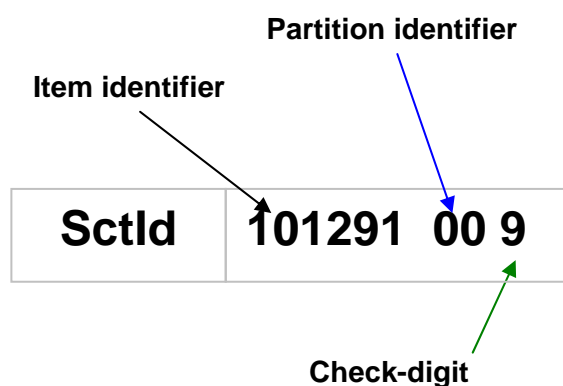
The SctId does not contain semantic information related to the meaning of a concept or term. It does however have a structure that is designed to allow different types of terminological components to be recognized. The nature of a component can also be derived from the table in which a component is distributed. However, the advantage of partitioning the SctId is that it avoids reuse of the same



identifier for a different type of component – thus avoiding ambiguity. This also allows the nature of the identifier to be recognized when stored in a record or transferred in a message.

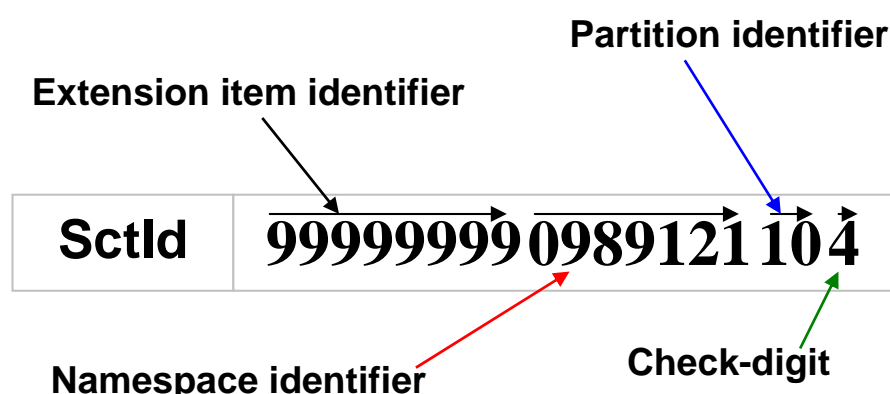
The rightmost digits in a decimal string representation of the SctId have the following defined roles:

- A single check-digit is used to validate the identifier.
- A two-digit partition-identifier, which:
  - Ensures that the identifier is unique in the scope of SNOMED CT; for example, the same identifier cannot be allocated to a Concept and to a Description.
  - Ensures that the identifier of a component in a valid Extension of SNOMED CT cannot be the same as the identifier of a component in the main body of SNOMED CT.



## 6.3 SctIds and Extensions

If the partition-identifier indicates that the SctId is part of an Extension the next seven-digits (from the right) are a namespace-identifier (see Figure 1). Namespace-identifiers are allocated to organizations which are authorized to issue Extensions. They enable unique SctIds to be issued by many organizations and allow each SctId to be traced to an authorized originating organization.



## 6.4 SctId Constraints

The constraints on value range for SctIds allow consistent string and integer representation of these values. Exclusion of negative values avoids potential formatting confusion in string representations.

The lower limit of 100,000 ensures the decimal string representation is at least six digits in length. This ensures that a SctId can be distinguished from:

- A Read Code, which is 5 characters in length, padded out with dots if necessary.
- A SNOMED ID, which always starts with a letter.

The upper limit of 18 digits ensures that any valid decimal string can be stored in either a signed or unsigned 64-bit integer.

## 6.5 Check-digit

The final (units) digit of the SctId is the check-digit. It is not envisaged that users routinely enter SctId values. However, the objective of the check-digit is to detect the commonest types of error that may occur due to typographical errors on those occasions where transcription or communication



mechanisms may introduce error. Examples may include high-level development such as creating or modifying protocols or pre-specified queries.

A SctId is checked by using the "Verhoeff check", which is a Dihedral  $D_5$  Check. This detects a higher proportion of common typographical errors than either the IBM or Modulus 11 check. Unlike the Modulus 11 check (used for the UK NHS number) it is effective on decimal strings longer than ten-digits. Furthermore its value can always be represented as a decimal digit without excluding any values.

See the *SNOMED Clinical Terms<sup>®</sup> Technical Reference Guide* for detailed information about the Verhoeff check-digit and sample program code.

## 6.6 Partition-identifier

The penultimate two-digits of the SctId (second and third from the right), are the partition-identifier.

The partition-identifier indicates the nature of the entity identified. This allows the identifier of a Description to be distinguished from the identifier of a Concept. It also allows SctIds issued centrally to components of the main body of SNOMED CT to be distinguished from components issued as parts of Extensions.

Identifiers of components in the main body of SNOMED CT have one of the following partition-identifier values:

00	A Concept
01	A Description
02	A Relationship

Identifiers of components in Extensions have one of the following partition-identifier values:

10	A Concept in an Extension
11	A Description in an Extension
12	A Relationship in an Extension



All other partition-identifier values are reserved for future use.

## 6.7 Extension namespaces

If the partition-identifier indicates that the SctId is part of an Extension the seven-digits immediately to the left of the partition-digit are a namespace-identifier.

Each organization that is authorized to issue Extensions is allocated a namespace-identifier. The authorized organization is permitted to assign any valid item identifier that ends with this string of digits.

SNOMED CT core release files include Namespace Concepts representing each of the allocated Namespace-identifiers. These Concepts have the following characteristics:

- They are direct subtypes of the Concept “Namespace Concept” which is a direct subtype of the Concept “Special Concept”.
- The Fully Specified Name has the form “Extension Namespace {nnnnnnn} (namespace concept) – where nnnnnnn is the seven digit Namespace-identifier.
- A Synonym associated with the Concept has the form “Extension Namespace nnnnnnn”
- Where appropriate further Synonyms may be included to identify the nature of the Extension and/or responsible organization. However, this information may not be made available for all Namespaces due to privacy constraints.

## 6.8 Item identifier digits

The remaining digits to the left of the partition-identifier (or in the case of Extensions, to the left of the namespace-identifier) are available to uniquely identify an individual entity within the specified partition. The same item identifier can be allocated in each partition and is rendered unique by the partition-identifier.

For components in the main body of SNOMED CT, item identifiers are usually issued in the arbitrary order in which components are added to SNOMED Clinical Terms. Due to management of the editing



process the sequence of issued item identifiers may be discontinuous and the order of identifiers should be regarded as meaningless.

## 6.9 Example Sctlds

The following SctId examples are based on the above rules and illustrate the range of possible item identifiers within each partition.

SctId	Partition identifier	Check digit	Notes
100005	00=Concept	5=OK	The Item identifier digits '1000' are the lowest permitted value thus this is the lowest SctId that can be allocated to a Concept.
100014	01=Description	4=OK	This is the lowest SctId that can be allocated to a Description.
100022	02=Relationship	2=OK	This is the lowest SctId that can be allocated to a Relationship.
101291009	00=Concept	9=OK	A valid SctId for a Concept.
1290023401015	01=Description	5=OK	A valid SctId for a Description.
9940000001029	02=Relationship	9=OK	A valid SctId for a Relationship.
10000001105	10=Extra-Concept	5=OK	A valid SctId for a Concept in an Extension in the 0000001 namespace.
10989121108	10=Extra-Concept	8=OK	A valid SctId for a Concept in an Extension in the 0989121 namespace.
1290989121103	10=Extra-Concept	3=OK	A valid SctId for a Concept in an Extension in the 0989121 namespace.
1290000001117	11=Extra-Description	7=OK	A valid SctId for a Description in an Extension in the 0000001 namespace.
9940000001126	12=Extra-Relationship	6=OK	A valid SctId for a Relationship in an Extension in the 0000001 namespace.
999999990989121104			The maximum valid SctId for a Concept in an Extension in the 0989121 namespace.