Migrating GP Primary Care Systems to SNOMED CT: A guidance document				
Programme	NPFIT	**Document Record ID Key**		
Sub-Prog / Project	Data Standards	NPFIT-FNT-TO-TOSCTI-0024		
Prog. Director	Ken Lunn	Status	Approved	
Owner	Keith Naylor	Version	2.0	
Author	John May	Version Date	August 2010	

Migrating GP Primary Care Systems to SNOMED CT: A guidance document

Amendment History:

Version	Date	Amendment History
0.1		First draft for comment
0.2	10 November 2009	Draft for formal review after inclusion of peer comments
1.0	21 December 2009	Document approved for discussion purposes
1.1	15 February 2010	Changes incorporated following discussions with GP System suppliers
1.2	June 2010	Changes incorporated following wider consultation – mainly to aid readability
1.3	July 2010	Changes made following formal review
1.4	July 2010	Changes made following further review
2.0	August 2010	Document approved for publication

Forecast Changes:

Anticipated Change	When
Update to Section 10 – Further issues to be considered.	

Reviewers:

This document must be reviewed by the following:

Name	Signature	Title / Responsibility	Date	Version
Jeremy Rogers		Principal Terminologist	15 July 10	1.2
Ian Green		Principal Terminologist	14 Dec 09	0.2
Steve Bentley		Senior Clinical Consultant	15 July 10	1.2
Denise Downs		Implementation and Outreach Lead	29 July 10	1.4
John Williams		Clinical Lead	15 July 10	1.4
Rae Long		Standards Consultancy Lead (Acting)	19 Nov 09	0.2

Approvals:

This document must be approved by the following:

Name	Signature	Title / Responsibility	Date	Version
Ian Arrowsmith		Chief Terminologist	29 July 10	1.4
Keith Naylor		Head of Implementation - DS&P	29 July 10	1.4

Distribution:

This document will be made available to clinical system suppliers who currently have systems which use Read2 or CTV3 as the native terminology.

The document will also be made available to other partner organisations as necessary.

Document Status:

This is a controlled document.

Whilst this document may be printed, the electronic version maintained in FileCM is the controlled copy. Any printed copies of the document are not controlled.

Related Documents:

These documents will provide additional information.

Ref no	Doc Reference Number	Title	Version
1	http://www.connectingforhealth.nhs.uk/factsandfiction/jargonbuster/acronyms	Glossary of Terms Consolidated.doc	

Glossary of Terms:

List any new terms created in this document. Mail the NPO Quality Manager to have these included in the master glossary above [1].

Term	Acronym	Definition

Contents

1 About this Document

1.1 Purpose

The purpose of this document is to provide guidance on a possible way for a Primary Care system that is currently based upon READ2 or CTV3 to evolve to a system that uses SNOMED CT as its core terminology.

This document is issued under the auspices of the UK Terminology Centre (UKTC) and is supported by the members of the UK Terminology Centre Implementation Forum (UKTC IF). The UKTC IF comprises representatives from the commercial IT supplier community, the NHS home countries, the UKTC and other interested parties.

The document is not intended to imply a strategy nor provide definitive implementation guidance for system migrations. Rather it suggests and explores one option of developing a GP system as a vehicle to expose and highlight the technical issues around the development and implementation of a GP system to become fully SNOMED CT enabled[1]. In addition the document does not address system or user interface design; any references to such are purely for illustrative purposes.

A number of GP Systems suppliers were consulted during the writing of this document and all support the methodology outlined as being feasible, albeit without the level of detail that would allow a full technical appraisal.

At some later point, it is expected that the content of this document will evolve into a firmer set of implementation guidelines. However these are outside the scope of the current document.

1.2 Audience

This document has been written primarily for the existing GP System Suppliers. However, it is equally useful to explain to a non-SNOMED CT audience how implementation could be achieved.

Whilst this document discusses some of basic SNOMED CT concepts; they are there simply to aid clarity, but it does mean readers are expected to have some technical understanding of how terminologies work within clinical systems. The document is therefore specifically not aimed at the general end user community.

1.3 Content

Over a series of sections, this document suggests and explores one option of developing a GP system as a vehicle to expose and highlight the technical issues around the development and implementation of a GP system to become fully SNOMED CT enabled. It does not aim to exhaustively list all functionalities that would become available at various steps along the pathway described; neither does it suggest that the implementation process is necessarily serial (e.g. Step A, then B, then C etc.).

[1] The term 'fully SNOMED CT enabled' is used in this document to describe a state where all the major benefits of SNOMED CT functionality are being realised within a system. This does not imply that everything that is possible with SNOMED CT has to be implemented.

For more detailed information, the reader should consult the SNOMED CT Technical Reference and Implementation Guides[2] . Less formal guidance can be obtained through the UK Terminology Implementation Forum (UKTC IF)[3] or the International Health Terminology Standards Development Organisation (IHTSDO) implementation special interest group[4].

Sections 2 and 3 cover the background and set the scene and vision for the future.

The next sections propose that the vision can be achieved by identifying the necessary steps that need to be taken and groups these into four phases:

- Phase Zero – A thin SNOMED CT veneer[5] over the legacy terminology
- Phase One – A thick legacy terminology veneer over SNOMED CT itself
- Phase Two – SNOMED CT with restricted functionality
- Phase Three – Full SNOMED CT functionality

Section 9 summarises the document.

It is important to point out that the content of this document has been reviewed and discussed in some detail directly with the GP System suppliers who have all individually confirmed its feasibility. It should also be noted that the GP System suppliers also agree that Phase One and some of Phase Two is achievable with current (2010) technology.

However the plan looks beyond this point and discusses future functionality; the guidance here being necessarily more generic with some elements of the technology discussed still to be developed and implemented in an operational system. This has been done deliberately in an attempt to provide a credible roadmap to achieve the technical vision of a fully SNOMED CT enabled system. It is also worth commenting that this functionality is already available in some applications outside the UK.

2 Background

Existing primary care systems are based on either the Read Codes V2 (READ2) or Clinical Terms Version 3 (CTV3).

SNOMED CT differs from READ2 and CTV3 in a number of significant ways which have to be taken into account right from the outset when considering changing terminologies. These differences may be categorised across the following broad dimensions:

Technical

(a) Physical presentation. For example: file sizes are much larger, character encoding is different (i.e. diacritic characters are allowed), code length is not fixed.

[2] See http://www.ihtsdo.org/publications/implementing-snomed-ct/
[3] The UKTC IF can be contacted via the NHS Data Standards Helpdesk data.standards@nhs.net
[4] The IHTSDO implementation special interest group can be contacted via info@ihtsdo.org
[5] The use of the term 'veneer' in this document is to indicate the layer of functionality specifically designed for the user interface, or to enable specific limited functionality

(b) Data model. For example the way that concepts are retired is completely different, the use of qualifiers and modifiers (like laterality, severity and urgency) are allowed.
(c) Knowledge content. For example:
    a. relationships are modelled allowing a polyhierarchical structure
    b. subsumption can be derived from the relationships
    c. there are multiple descriptions/terms for a concept (fully specified name, preferred term and any number of synonyms)
    d. There are far more concepts (by a factor of more than 10) [6]

Representational

(d) Relationships between concepts cannot be inferred from their identifiers. For example, in Read 2 the concept **C10..** equates to **Diabetes mellitus** and any concept that begins **C10** must be a form of **Diabetes mellitus**; i.e. **C10E.** equates to **Type 1 diabetes mellitus**. However in SNOMED CT **73211009** equates to **diabetes mellitus** and **46635009** equates to **diabetes mellitus type 1** – the concept identifiers are simply random
(e) Relationships are modelled and are explicitly stated**.** In other words, "*D* ***is a*** *B*" (*B subsumes D*, or *D is-subsumed-by B*) which means that concept *D* is a specialisation or more detailed type of concept *B*, and concept *B* is a generalisation of concept *D*. For instance, a "fruit" is a generalisation of "apple", "orange", "mango" and many others. One can say that an *apple* ***is a*** fruit.
(f) The content is polyhierarchical, meaning that a concept can have more than one logical parent. For example, in SNOMED CT: **diabetes mellitus** *IS_A* (type of) **disorder of the pancreas** and also *IS_A* (type of) **metabolic disease**

All of these factors will impact every existing GP clinical system, to a greater or lesser extent. To illustrate this: many, if not all existing primary care systems have been engineered such that they 'hard code' particular assumptions about how each of the above dimensions are managed in either READ2 or CTV3. For example, because READ2 represents its subsumption hierarchy within the concept code, subsumption queries may have been implemented as string comparison operations[7]. The migration of such systems to SNOMED CT as the underlying terminology requires, at a minimum, that the relevant software code is generalised[8]. This may be both costly and introduce business risk if significant amounts of application code must be re-factored to achieve full generalisation.

[6] A related point is that the logically computable foundation on which SNOMED CT is based will increasingly lead to a mismatch between the end users expectation of seeing clinical concepts classified heuristically and SNOMED CTs formal classification of them. While it can be argued that *all* terminologies/classifications at sometime return results at odds with user expectation, where such behaviour in SNOMED CT has resulted from logical inference, it is deliberate and will not be 'fixed' within SNOMED CT itself. Additional layers of knowledge may be required around the core SNOMED CT content.

[7] E.g. IF INSTR(Journal.Entry.Code,'C10F') =1 THEN …. /* Patient has diabetes */

[8] The reason for generalising the code is to recognise the fact that SNOMED CT has multiple and dynamic hierarchies, as opposed to the single static hierarchy of READ2. Hard-coding to a specific node may produce unexpected results over time.

A challenge for achieving SNOMED CT uptake, therefore, is how to either reduce - or at least stage - the cost of this migration: to provide the mechanism for change whilst at the same time minimising or staging the risk associated with that change.

3 Starting the process

Existing systems are likely to take different routes to becoming fully SNOMED CT enabled. However, it is considered highly unlikely that this will ever be done in one step due to the volume of change this entails; each supplier will need to find an acceptable and practical evolutionary route. The only exception to this would be where a System Supplier chose to develop a completely new SNOMED CT system from the ground up *and* where there was no requirement for any older existing product(s) to interoperate with the new 'green field' system.

'Brown field' migration of established systems is, therefore, expected to be more common. High level analysis of the differences between SNOMED CT and the legacy terminologies suggests a possible methodology within which to consider staging the effort and associated business risk required during such migrations from a legacy terminology to SNOMED CT. Phase Zero has been included to indicate what is believed to be the current status quo; i.e. the ability to be able to send and receive defined messages containing SNOMED CT thereby meeting the existing requirements for users and the NHS.

Broadly, the key to this suggested approach is in phase 1 for suppliers to focus effort initially on implementing, in full, the necessary physical and data model changes required by SNOMED CT, whilst simultaneously insulating themselves and end users almost entirely from any knowledge content changes that would otherwise normally arise immediately.

It is recognised that this proposal commits the system supplier to making considerable changes to their system with no immediate visible benefit to the user. All the GP System suppliers who contributed to this document agreed that for a system to be able to utilise the benefits of SNOMED CT, then this is the logical first step.

Subsequent to installing the new technical platform it would be possible to incrementally reveal and adopt the new knowledge content available through SNOMED CT and, thereby start to realise the benefits of SNOMED CT.

The individual steps outlined below are not intended to be necessarily implemented serially. However, they may be logically grouped together into a number of implementation phases:

- Phase Zero – A thin SNOMED CT veneer over the legacy terminology
    - **There may be some use of SNOMED CT but this is restricted to only that required to fulfill basic interoperability requirements**.
    - Any system in this category is classified as **SNOMED CT non-compliant**
    - This is the current state of most GP Systems in live use and is therefore considered to be the starting position in terms of this document.

- Phase One – A thick legacy terminology veneer over SNOMED CT itself
  - **The legacy terminology veneer here is to make the system look and function to the user as if nothing has changed**
  - A system in this category would be classified as **Level 1 SNOMED CT compliant**
- Phase Two – SNOMED CT with restricted functionality
  - **The legacy terminology veneer begins to be removed to reveal some functionality supported by SNOMED CT and not by the legacy terminology**
  - A system in this category is classified as **Level 2 SNOMED CT compliant**
- Phase Three – Full SNOMED CT functionality
  - The legacy terminology veneer has been removed to reveal SNOMED CT to the user
  - A system in this category is classified as being **Fully (or Level 3) SNOMED CT compliant**

These phases could also be viewed as indicating levels of maturity along the pathway from the current state to the point where the system, and starts to provide functionality that takes advantage of the features of the terminology.

From a user perspective, adopting this phased approach would initially (phase 1) have no visible impact as the 'look and feel' of the system would remain the same. In phase 2, the benefits of changing to SNOMED CT would begin to be manifest as more detailed clinical expressions become possible and new functionality is added, such as fully coded discharge summaries (rather than simply free text) and better interoperation with other organisations. In the final phase, the full functionality of SNOMED CT will be available to the user, for example full clinical decision support and interoperability between systems (both primary and secondary care). Throughout the process it should be possible to demonstrate improvements in patient care – which is everyone's ultimate goal.

In addition to making the changeover as easy as possible for the end users, the phases **could** be aligned with a compliance model to acknowledge the achievement of significant accreditation milestones. Dependencies between the three distinct phases toward compliance (i.e. ***not*** Phase Zero) mean that they necessarily follow each other. It is therefore possible to consider that accreditation could be against each phase rather than each step.

The phases are explained in further detail below.

4 Phase Zero – A SNOMED CT veneer over a legacy terminology

Some, if not all GP System Suppliers have already started to implement some SNOMED CT functionality. This limited functionality is presented to the users as a

thin veneer of SNOMED CT identifiers over a system that fundamentally still natively[9] uses READ2/CTV3 information and data models.

It is thought the systems that have adopted this approach have done so by utilising various cross-mapping tables; some of which have been published by the UKTC (but for different use cases) with others being hand-crafted by the suppliers themselves. This latter situation is raising concerns that cross-maps are being developed in an unconstrained way, which will inevitably lead to data anomalies in later phases - this is because significant maintenance effort will be required as updates of the terminology are released every 6 months.

An example of this phenomenon is some of the work to co-opt a locally specified and locally enhanced subset of the National Interim Clinical Imaging Procedure (NICIP) catalogue to support radiology order communications in primary care. This is where the maps from NICIP Short Codes to READ codes have been done without reference to either the SNOMED CT maps explicitly provided with NICIP or the UKTC SNOMED CT to READ2 maps.

There is also the issue of clinical validation of any cross-mapping tables – but this is out of the scope of this document. However, this is currently being addressed by a separate piece of work being undertaken by the UKTC.

In the short-term, applying ever thicker veneers of SNOMED CT over a legacy terminology may seem to be a pragmatic solution. However, this approach is not commercially sustainable to the NHS in perpetuity[10] and can never offer the level of functionality that will be available in a native SNOMED CT solution. Because of this, a solution such as this will not constitute an approved SNOMED CT implementation as described in the NHS GP Systems of Choice compliance levels[11].

GP System Suppliers whose systems already meet this level of maturity are strongly encouraged to move directly to Phase One.

5 Phase One - A legacy terminology veneer over SNOMED CT

Whilst minimising and staging change is seen as a logical commercial step, UKTC would prefer to support a ‘thicker veneer’ through which SNOMED CT gains the surface appearance of a legacy terminology rather than vice versa. In order to assist development of such ‘thick veneer’ solutions, the UKTC proposes to provide system suppliers with ‘views’ on SNOMED CT that explicitly transform its content so that it is possible to make a system appear, to the end user, to be the same as either READ2 or CTV3. These views will be generated directly from and delivered within SNOMED CT native mechanisms and would comprise:

[9] The use of the term ‘native READ2’ or ‘native SNOMED CT’ etc. is a way of describing the core terminology around which the application is written
[10] The cost of developing, maintaining and ensuring synchronicity between three separate terminologies impacts all areas of the NHS, not just the UKTC
[11] http://www.connectingforhealth.nhs.uk/resources/archive/gpsococtupdate.pdf

1. a SNOMED CT concept subset/RefSet listing all the concepts in SNOMED CT that are equivalent to a concept also in READ2/CTV3,
2. a SNOMED CT description subset/RefSet listing all the descriptions for the SNOMED CT concepts in (1) that also appear for the same concept in READ2/CTV3 (and, which is the preferred description for display, natively in READ2/CTV3),
3. a SNOMED CT navigational subset/RefSet listing all the SNOMED CT-native parent-child relationships that exist between some pair of SNOMED CT concepts in (1) and that also hold for the corresponding concepts in READ2/CTV3, and also adding any additional relationships that are true in READ2/CTV3 but that are not found in SNOMED CT. This would be used so that concept subsumption test (e.g. to drive QOF queries) would return the same answers as if the clinical concepts concerned were being analysed using the legacy terminology concept hierarchy,
4. concept cross-map tables to enable bidirectional conversion between SNOMED CT and READ2/CTV3 concept identifiers.

Note

It is possible that suppliers may be using other subsidiary artefacts normally included in either the READ2 or CTV3 releases[12]. Therefore these would also have to be re-expressed within the SNOMED CT model in order that information could be provided with exactly the same knowledge content but via the SNOMED CT data model.

These artefacts need to be validated and maintained, however the UKTC can only be responsible for those it creates. The process for validating and maintaining those artefacts that reside outside the UKTC domain is less clear, for example where a GP Supplier creates their own.

The technical platform required to implement this solution – to load SNOMED CT in its entirety but to retain the look and feel of READ2 or CTV3 – would be more complex than the 'thin veneer' option. However, it would enable an incremental migration toward full SNOMED CT implementation and benefits realisation; through the following implementation steps:

A. Extending text-to-concept searching to include all SNOMED CT descriptions valid for concepts in the transform concept subset (this would theoretically provide more accurate searching).

   As discussed above, SNOMED CT has the ability to hold many synonymous terms for a single concept; this then makes it easier for the user to retrieve the most appropriate code for what they are trying to say. This would also introduce the UTF-8 standard, which allows for storage, search and display of diacritic characters, such as in **ménière's disease**

[12] E.g. the codes with values list, or PBCL

B. Extending text-to-concept searching to include all SNOMED CT descriptions valid for concepts in the transform concept subset OR for any of their descendents (this would theoretically provide more accurate searching and coding).

   For example, If READ2 is searched for the expression '**Houssay syndrome**' nothing will be returned. If the same search is done against SNOMED CT, it returns **123763000 Houssay's syndrome (disorder)**, but this would not be in the set of concepts specified in any READ2 legacy veneer subset. However, one of its ancestors is: **73211009 Diabetes mellitus (disorder).** So, whereas the search expression run natively in the legacy READ2 terminology *should* draw a blank, if the veneer is pulled back and run against SNOMED, it can return **73211009** (or **C10..** if READ2 codes are still being displayed) as being the best available match in Read2 for the original search expression.

C. Appropriate filtering of candidate matches returned from a 'search for select'.

   This item is intended to include filtering of candidate string match result sets by RefSet membership, as well as possibly by the semantic relationship and metadata properties involved. Therefore this item might actually unpack into three or more discrete sub-steps, although they would not have to be applied in any particular order:

   1. Filter candidate matches by metadata: such as concept status (e.g. if the concept is currently allowed to be used); identifier namespace (e.g. if the concept is valid for use only in England); language (e.g. Spanish) etc.
   2. Filter candidate matches by membership of a RefSet (a RefSet is a defined collection or list of concepts). This further divides into enumerated or intentionally defined RefSet.
   3. Filter candidate matches by semantic properties; i.e. whether a concept has particular kinds of semantic relationships to particular kinds of other concepts, e.g. it is a type of clinical finding.

By the time all of this work has been completed:

- The system will be using SNOMED CT natively.
- Clinical records will be recorded using SNOMED CT but will only allow those within the transform RefSet.
- There should be no impact on interoperability, messaging and reporting.

At the end of this phase, a system would be eligible to be classified as **Level 1 SNOMED CT Compliant.**

6 Phase Two – SNOMED CT with limited functionality

Building upon work so far carried out in Phase One, the next stages would be:

D. Extending coding choices to include SNOMED CT concepts that are subsumed by those concepts in the transform concept subset. This is similar to step B, but in this case any SNOMED CT term that is child of a concept in the subset can be selected (i.e. not just those that have equivalent Read2/CTV3 parents).

E. Reporting functionality will now need to take account of the changes to the coding structure. As soon as SNOMED CT concepts are used that are not mapped back to READ2/CTV3, then it will no longer be possible to rely on the corresponding legacy concept code alone to assert relationships (as described earlier).

F. Extending coding choices to include all SNOMED CT concepts (subject to application of further filtering of candidate matches to exclude contextually inappropriate selections – for example only those concepts that appear in a specific hierarchy(ies)). In effect this is opening up SNOMED CT completely for concept selection. However the concepts that would be made available would depend upon the context of the search required. For example if the user is trying to enter a diagnosis, then it would be logical to show only those concepts that are in the SNOMED CT clinical findings hierarchy.

G. Displaying the concepts in the transform subset as a polyhierarchy for the purposes of hierarchy navigation in a code-picking context. Until this point, the subset has been constructed and displayed to resemble Read2; i.e. as a single hierarchy. This step removes that artificial layer and thus exposes the polyhierarchical nature of SNOMED CT to the user thus allowing much easier navigation and searching of the concepts. It is worth noting that users could be given the option to switch back and forth between the two hierarchy views for a period of time.

H. Analysing the concepts in the transform subset as a polyhierarchy for the purposes of decision support and/or reporting (e.g. QOF). This is where the power and benefits of SNOMED CT's polyhierarchical structure over the monohierarchy of READ2 becomes truly visible.

I. Validate incoming expressions received from external systems against the SNOMED CT concept model constraints[13], i.e. the ability to machine process incoming expressions. In effect the system would have to internally store and manipulate SNOMED CT post-coordinated expressions coming from outside, even if users are not able to create them. A use case for doing this is in order to be able to validate incoming expressions and store them.

J. Render incoming post-coordinated expression in human readable form on screen. SNOMED CT has the ability to link concepts together to allow more complex statements to be made – hence post-coordination. At this point, the

[13] using the IHTSDO Machine Readable Concept Model

system will be able to receive such post-coordinated expressions and display them to the user in a form that is understandable.

K. Offering limited 'simple' post-coordination choices, e.g. laterality, urgency, priority etc[14] [15]. In effect this 'decorative' post-coordination will only allow detail that would normally be disregarded for practical classification purposes. Ideally, such expressions should still only be constructed within the confines of what is an 'allowable post-coordination' (within the published concept model). So this step therefore also implies implementing the optional qualifier rows and/or the IHTSDO/UKTC Machine Readable Concept Model (MRCM).

   The reason for introducing this level of post-coordination is to introduce end users to the concept of how to create expressions within SNOMED CT rather than relying on the system to do it for them. This is in preparation for the time when users will be allowed to produce much more complex expressions.

By the time all this work has been completed:

- The system will be using SNOMED CT as the native terminology.
- All historical clinical records will have been converted to SNOMED CT and new entries will only be recorded using SNOMED CT.
- If full functionality is to be maintained, there is likely to be some impact on interoperability, messaging and reporting, however this is believed to be minimal.
- Users will have been exposed to some of the functionality available within SNOMED CT in a controlled and understandable manner. It is thought that by this time, systems will look and feel quite different, which means that there will undoubtedly be a need for significant user training.

At the end of this phase, a system would be eligible to be classified as **Level 2 SNOMED CT Compliant.**

7 Phase Three – Full SNOMED CT functionality

Building further upon the previous work, to the point where the full functionality of SNOMED CT can be utilised within the system:

L. Wider use of native 'non-decorative' post-coordination. For example, <finding> suspected; i.e. Swine flu suspected.

[14] Note that this is different to holding the data in the information model and creating post-coordinated expressions on the fly.
[15] Note that draft guidance on compositional grammar has been published by IHTSDO. Whilst primarily aimed at HL7 messaging, the document is a useful read for those interested in post-coordination.

Such post-coordination expression building MUST be within the confines of what the MRCM and/or optional qualifier constraints says is allowable, whilst querying and reporting MUST also be done with appropriate reference to the full semantics of the post-coordinated expression[16]. Simply put, post-coordination must fit within the current rules and guidance.

a. Step L but where querying can be satisfied without recourse to a full description logic classification process (a brief introduction to Description Logic can be found at appendix 1).

b. Step L but where querying is satisfied only by use of a description logic classifier.

M. From step L there is a need for more advanced term selection (and expression construction) interfaces. For example, finding the right body site for a burn lesion probably should not be done by a picking list match against text.

In parallel to steps L and M is whether the query specifications, as well as new coded data items, can be post-coordinated.

Those steps (A - M) outlined so far represent incremental adoption firstly of SNOMED CT's richer lexical corpus, then of SNOMED CT's different static (enumerated, pre-coordinated) subsumption hierarchy, then finally of SNOMED CT's dynamic (post-coordinated) subsumption hierarchy.

Other components of SNOMED CT knowledge content – particularly RefSets and Crossmaps - exist somewhat in parallel to this central incremental interaction with the core sct_descriptions and sct_relationships tables. Their adoption opens up other functional enhancements of their own and leads on to additional migration steps:

N. Query definition software using RefSets to store, or filter, certain results.

O. Query specifications written in terms of intentional RefSet definitions (with or without post-coordinated concept expressions).

P. Implementation of cross-maps to external schemes using the native SNOMED CT information model for encoding crossmaps.

By the end of this technical migration pathway, the fully compliant SNOMED CT platform will necessarily include several new technologies including: a post-coordinated expression building GUI (graphical user interface), an expression constraint rule set (e.g. the MRCM) and associated expression validation engine, a Description Logic classifier and a post-coordinated expression storage solution. Full realisation of the benefits of SNOMED CT implementation depend on also adding value to the SNOMED CT encoded record by reference to other resources that are primarily encoded by reference to SNOMED CT expressions, for example Map of

[16] The UKTC Reporting project (end date: May 2010) is attempting to determine whether it is possible to be more precise about the implementation implications of this shift.

Medicine or British National Formulary. These external resources may themselves also be encoded within the SNOMED CT information model publishing paradigm.

At the end of this phase, the system would be eligible to be classified as **Fully (or Level 3) SNOMED CT Compliant.**

8 Core functionality

The steps outlined above set out a framework within which to consider possible pathways for the incremental adoption of SNOMED CT's knowledge content. Such pathways would seek to stage and manage any system behaviour changes that will arise where SNOMED CT's characterisation of the clinical domain space differs from that represented in the knowledge content of a legacy terminology.

These steps, however, presuppose and are variously dependent on the prior or parallel implementation of a supporting technical SNOMED CT platform. This platform shall include a range of minimum required functionalities that cope with new surface properties of SNOMED CT's knowledge content and, in particular, with the mechanisms for tracking changes in that content. A non-exhaustive list of these functionalities includes:

1. Screen display, database storage, indexation and string matching across the UTF-8 character set [required from step A].
2. Application of filters to exclude certain SNOMED CT concepts from display and/or selection (e.g. non-human subset) [required from step B].
3. Management of the consequences of SNOMED CT concept retirement with respect to historically encoded data or reporting queries [required from step D].
4. Expression transformation (e.g. generation of normal forms) and validation (e.g. against the SNOMED CT Machine Readable Concept Model) [required from step H].
5. Dynamic classification [required from step H].

Note that items 3 and 4 are likely to be major pieces of work.

A major consideration here is that throughout the evolution of systems to become SNOMED CT compliant, consideration needs to be given to the preservation of functionality from the end user perspective. This is not saying that the functionality must be identical in every way; indeed it is highly likely that there will be changes at the user interface. It is envisaged that such changes will make the system easier for the end user to operate.

9 Summary

In summary, having discussed the content of this document with five of the six UK GP Systems Suppliers, it is felt that the true barrier to SNOMED CT implementation is not primarily to do with any deep rooted uncertainties over how a SNOMED CT implementation should work internally, but the much trickier problem of how to swap

out any one terminology for another whilst satisfying all the clinical safety processes at the same time as also maintaining business continuity for reporting during the switchover and after.

Many of the steps that, pragmatically, may need to be taken to mitigate the business continuity issues (e.g. accept some degradation of service; invest in training) are not within the influence of UKTC or GP System Suppliers.

Essentially, with the introduction of Commissioning Bodies and GP Consortia, there are increasing requirements for evidence based practice across all care settings. In the past, READ2 and CTV3 were adequate to meet the requirements of the Primary Care sector computerised Electronic Health Record. However, now that other health care sectors are beginning to deploy EHR's, the need for cross care setting interoperability is becoming evident. This can only be achieved effectively and efficiently if there is a common lingua franca – SNOMED CT.

This document demonstrates a credible and feasible roadmap for the GP System suppliers to move their current systems to one that is fully SNOMED CT compliant.

Should any of the NHS home countries choose to turn this feasibility into some form of reality then the next step would be to develop these proposals into a specification against which systems could then be accredited. The UKTC IF could assist in this if required.

10 Appendix 1 – an introduction to Description Logic

Description Logic(s) (DL) is a family of formalisms that are used to represent knowledge in domains like biomedical sciences. DL uses concepts, properties (roles) and individuals (instances) to represent the domain. A concept is a kind of a thing in the given domain; e.g. Bacterium, Organism, Pneumonia, Lung, Respiratory_System_Part, etc. A 'property' is a 'character' of a 'concept'; e.g. Pneumonia can have properties 'causative_agent' or 'finding_site'. Since SNOMED CT uses DL, SNOMED CT concepts and attributes correspond to DL concepts and properties.

One salient feature of DL is the use of a 'reasoning service', commonly referred to as the 'reasoner' or 'classifier'. The DL reasoner allows for implicit inferences to be made from explicitly stated knowledge in the knowledge base. Such implicit inferences allow for 'classification', where concepts in the knowledge base are organised on the basis of 'parent - child' relationships (also known as subsumption relationships). For example; given the following explicit statements in the knowledge base, where the symbol |_ represents an 'is a type of' relationship:

Organism

|__ Bacteria

Respiratory_System_Part

|__ Lung

Disease

|__ Disease_of_Respiratory_System

|__ Disease_caused_by_Organism

|__ Bacterial_Pneumonia

|__ Lung_Disease

The following description for Bacterial_Pneumonia can be created:

Bacterial_Pneumonia =

Disease

has_finding_site = Lung

has_causative_agent = Bacteria

The rest of the disease is then defined as follows:

Lung_Disease =

Disease

has_finding_site = Lung

Disease_of_Respiratory_System =

Disease

has_finding_site = Respiratory_System_Part

Disease_caused_by_Organism =

Disease

has_causative_agent = Organism

Note that at this point, it has only explicitly stated that 'Bacterial_Pneumonia' is a 'Disease'. The explicit statements in the definitions of the rest of the disease do not relate to 'Bacterial_Pneumonia'. If the DL reasoner was now run, it would now classify the disease hierarchy above as follows:

Disease

|__ Disease_of_Respiratory_System

|__ Lung_Disease

|__ Bacterial_Pneumonia

|__ Disease_caused_by_Organism

|__ Bacterial_Pneumonia

The DL reasoner generates these new hierarchical (IS_A) relationships based on the implicit knowledge contained in the explicit statements in the knowledge base. The 'IS A' relationships in SNOMED CT are generated by the College of American Pathologists (CAP), using a DL reasoner.