
Matt Cordell

**Results of Analysis
of SNOMED CT
Extensions**

Author: Matt Cordell

Version: 27

Date: 12-Apr-2017 06:08

Table of Contents

| | | |
|----------|---|-----------|
| 1 | Introduction | 4 |
| 1.1 | A raw analysis of the active concepts within an extension. | 5 |
| 1.2 | Ratio of active to inactive concepts | 6 |
| 1.3 | Extension changes against International Concept IDs | 6 |
| 2 | Extension Concepts | 8 |
| 2.1 | The type of content by hierarchy | 12 |
| 2.1.1 | Clinical finding | 14 |
| 2.1.2 | Procedure | 16 |
| 2.1.3 | Special concept | 17 |
| 2.1.4 | Situation with explicit context | 19 |
| 2.1.5 | Observable entity | 20 |
| 2.1.6 | Event | 21 |
| 2.1.7 | Qualifier value | 21 |
| 2.1.8 | Record artifact | 23 |
| 2.1.9 | Social context | 24 |
| 2.1.10 | Substance | 25 |
| 2.1.11 | Body structure | 26 |
| 2.1.12 | Staging and scales | 27 |
| 2.1.13 | Pharmaceutical / biologic product | 28 |
| 2.1.14 | Organism | 29 |
| 2.1.15 | Environment or geographical location | 30 |
| 2.1.16 | Specimen | 30 |
| 2.1.17 | Physical object | 31 |
| 2.1.18 | Physical force | 31 |
| 3 | Extension Descriptions | 32 |
| 3.1 | Extension Changes to Core Descriptions | 33 |
| 4 | Relationship Extensions | 34 |
| 4.1 | Core relationships modified within an Extension | 35 |

| | | |
|-----|--|----|
| 4.2 | Types of Relationships Modified | 36 |
| 5 | Additional Observations | 41 |
| 6 | Extensions vs Editions | 42 |
| 6.1 | File naming | 42 |
| 6.2 | Directory structure | 42 |
| 6.3 | Specific file inclusions | 43 |
| 6.4 | Miscellaneous QA issues | 44 |
| 7 | Conclusion | 45 |
| 8 | Endnotes | 46 |

1 Introduction

In January 2017, the Content Managers Advisory Group (CMAG) initiated an action to conduct a survey of the national extensions that were available, as to date there was limited information available - to other members and presumably SNOMED International. The responses of this survey are available [here](#). The results shows a variety of extensions are produced, ranging from subset/refset, language translations and clinical content development. Subsequent activities investigating collaboration on subset may be pursued, but the CMAG was also interested in exactly what clinical content was in each extension, and how this might be shared. There should be very little clinical content that is exclusive to any country, and if it has been developed by one extension - it's likely globally relevant, and sharing it can reduce duplicated effort and maintenance burden of extension builders.

The results described here are a result of a combination of objective metrics (size of content), crude identification of duplicated effort, and finally some incidental quality observations. Further analysis of the content is still underway using description logic techniques, the results of which will be made available separately, at a later date.

SQL snippets are included in the document for future reference by author, but will unlikely be useful to public readership.

A summary of the results is available in the [conclusion section at the end of this paper](#).

The cooperation of all Members is appreciated, and whilst all effort has been made to represent the extensions accurately, any inaccuracies are accidental.

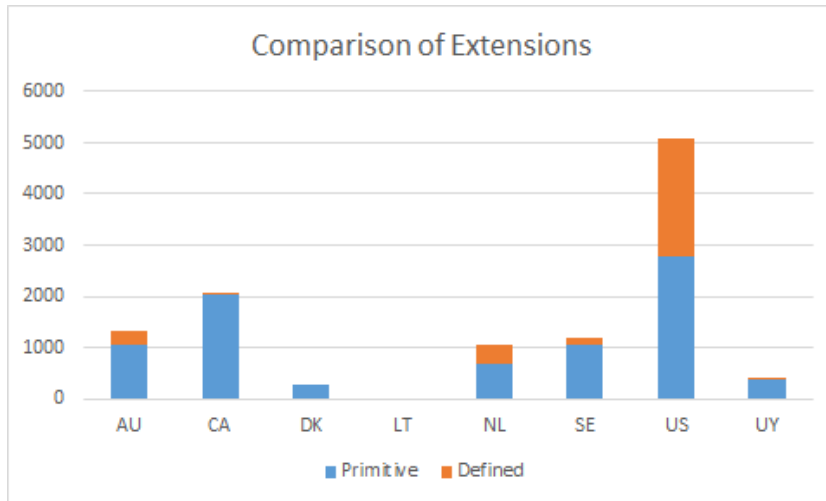
Summary of Extensions

14 NRC responded to the survey, with **9** indicating they created clinical content extensions.

The Australian Edition also includes it's national drug extension, which has been excluded from this round of analysis (as no other extension appeared to include such content)

All extensions were based upon the July 2016 international release except one. This exception may produce some anomalies, but they are limited to the extension.

1.1 A raw analysis of the active concepts within an extension.



Note: UK extension (~30 thousand concepts) is about 6 times larger than US (5000) and has been excluded from the graph as an outlier. Raw values are shown to the right.

| NRC | Size | Primitive | | Defined | |
|-----|-------|-----------|--------|---------|-------|
| AU | 1339 | 1056 | 78.9% | 283 | 21.1% |
| CA | 2040 | 2030 | 99.5% | 10 | 0.5% |
| DK | 290 | 290 | 100.0% | 0 | 0.0% |
| LT | 0 | 0 | N/A | 0 | N/A |
| NL | 1041 | 686 | 65.9% | 355 | 34.1% |
| SE | 1175 | 1060 | 90.2% | 115 | 9.8% |
| UK | 29793 | 29793 | 100.0% | 0 | 0.0% |
| US | 5100 | 2764 | 54.2% | 2336 | 45.8% |
| UY | 415 | 371 | 89.4% | 44 | 10.6% |

1.2 Ratio of active to inactive concepts

| NRC | Proportion currently active |
|--|-----------------------------|
| SNOMED CT Netherlands NRC maintained module | 95.3% |
| US National Library of Medicine maintained module | 93.3% |
| módulo de la extensión de Uruguay | 95.6% |
| Canada Health Infoway English module | 68.3% |
| SNOMED CT Sweden NRC maintained module | 99.6% |
| Australian common model component extension | 33.3% |
| Danish module | 88.4% |
| SNOMED Clinical Terms Australian extension | 97.0% |
| SNOMED CT United Kingdom clinical extension module | 32.9% |

All analysis was only performed on active content.

1.3 Extension changes against International Concept IDs

A total of 40 core concepts have been modified by extensions in some way.

Two were retired by an extension

- 384612007| pT4a: Tumor directly invades other organs or structures (colon/rectum) (finding) |
- 384613002| pT4b: Tumor penetrates visceral peritoneum (colon/rectum) (finding) |
- (A third concept was retired, but later reactivated)

One concept had a change to definition status (marked *Defined*) by an extension

- 399733007| Excision of retroperitoneal lymph node (procedure) |

Results of Analysis of SNOMED CT Extensions

Eight of these appear to be an attempt to address issues within the module assignment in the international release. (i.e. Concept inactivated on a different module to what they were created. metadata vs core).

The remainder are simply changes to moduleId, and either represent content promotion from an extension to the International. Or a possible error.

```
select id, count(*) from X_Concepts
where id in
(246089008, 246221002, 260670006, 263512003, 263513008, 447564002, 44960
9005, 700043003, 11000119105, 41000179103, 441000119109, 601000119109, 1
111000119100, 1561000119105, 4181000179103, 4191000179101, 42010001791
04, 4211000179102, 4221000179107, 4231000179109, 4241000179101, 4251000
179103, 4261000179100, 4271000179106, 4281000179108, 4301000179109, 431
1000179106, 4321000179101, 4331000179104, 4341000179107, 4351000179105
,
5461000179100, 5471000179106, 5481000179108, 5491000179105, 5531000179
105)
and moduleId != 161771000036108
group by id
having count(distinct moduleId) > 1
```

2 Extension Concepts

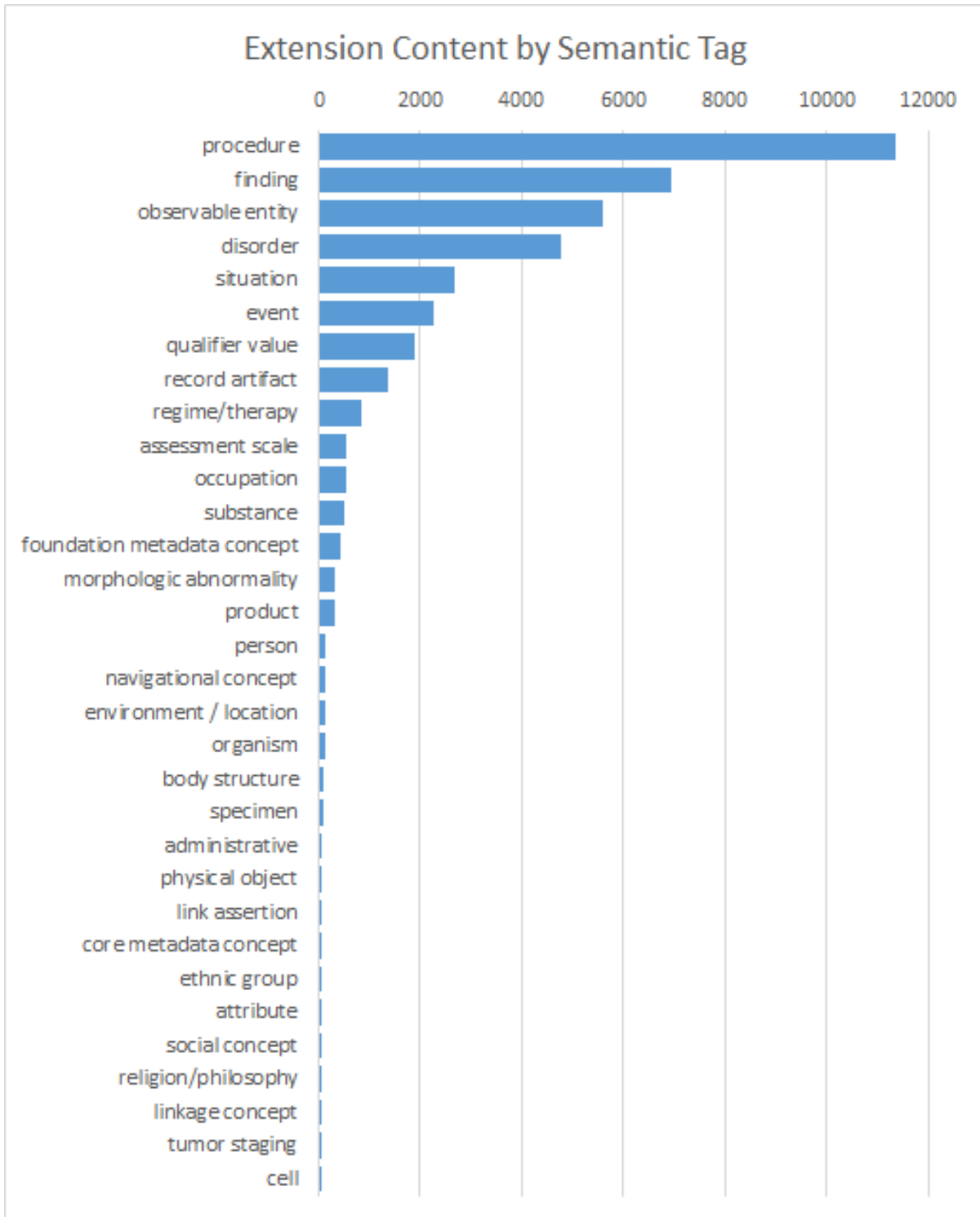
There appears to be around 52 unique semantic tags across the extension content. many of these are attributable to translations. **Not all extensions provide en FSNs for extension content¹**, semantic tags were manually translated and merged.

After normalisation, this comes to 32 semantic tags. The distribution of content is shown below.

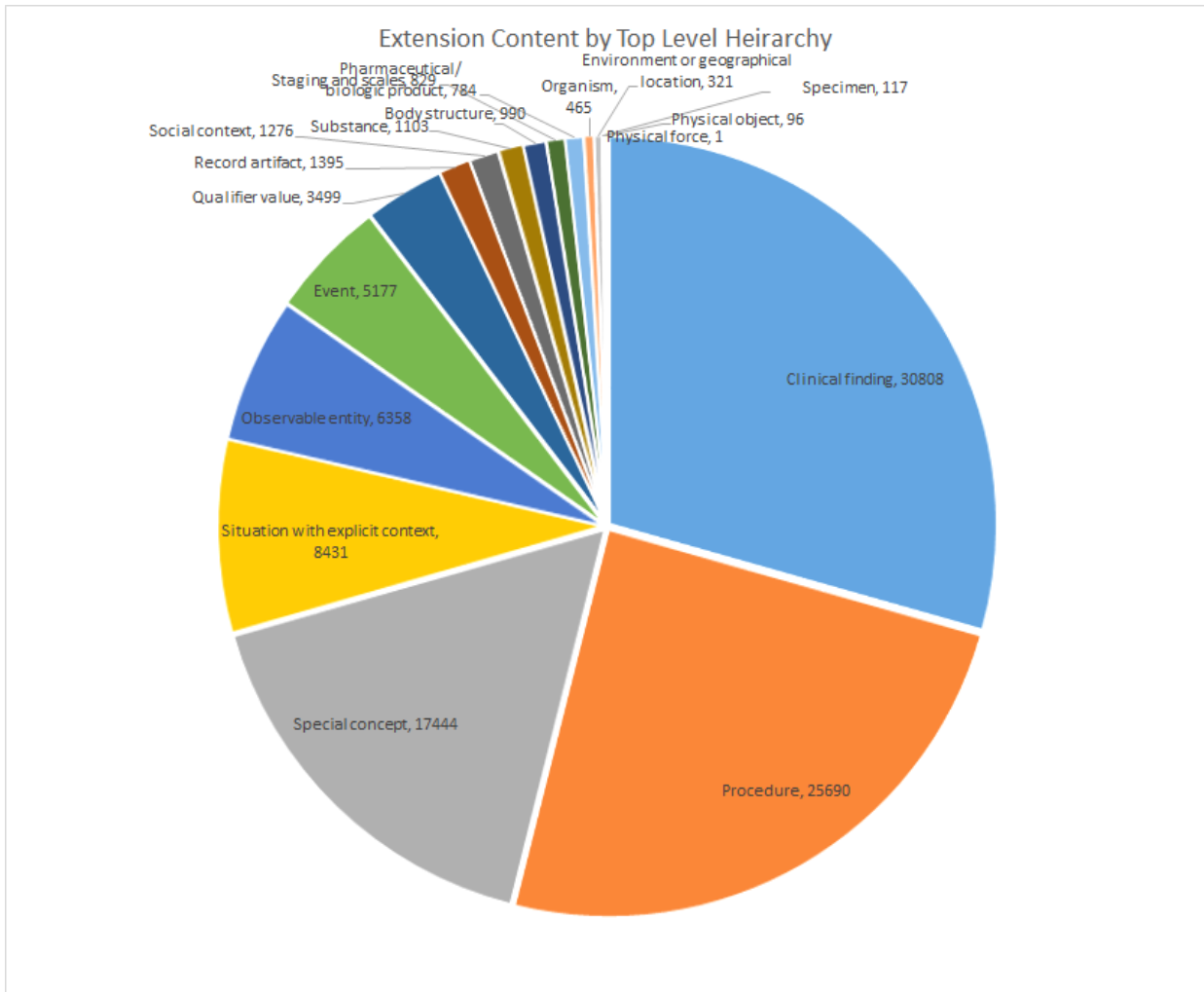
| Semantic Tag | Count |
|-----------------------------|-------|
| procedure | 11372 |
| finding | 6925 |
| observable entity | 5609 |
| disorder | 4783 |
| situation | 2677 |
| event | 2251 |
| qualifier value | 1903 |
| record artifact | 1361 |
| regime/therapy | 846 |
| assessment scale | 549 |
| occupation | 543 |
| substance | 503 |
| foundation metadata concept | 414 |
| morphologic abnormality | 321 |
| product | 320 |

Results of Analysis of SNOMED CT Extensions

| | |
|------------------------|-----|
| person | 134 |
| navigational concept | 132 |
| environment / location | 114 |
| organism | 114 |
| body structure | 106 |
| specimen | 95 |
| administrative | 66 |
| physical object | 59 |
| link assertion | 22 |
| core metadata concept | 19 |
| ethnic group | 16 |
| attribute | 7 |
| social concept | 5 |
| religion/philosophy | 4 |
| linkage concept | 3 |
| tumor staging | 2 |
| cell | 1 |



Results of Analysis of SNOMED CT Extensions



9 Modules are in use across the extensions.

| ModuleId | FSN | Country |
|-----------------|---|---------|
| 11000146104 | SNOMED CT Netherlands NRC maintained module | NL |
| 731000124108 | US National Library of Medicine maintained module | US |
| 5631000179106 | módulo de la extensión de Uruguay | UY |
| 20621000087109 | Canada Health Infoway English module | CA |
| 45991000052106 | SNOMED CT Sweden NRC maintained module | SE |
| 161771000036108 | Australian common model component extension | AU |

| ModuleId | FSN | Country |
|--------------------|--|---------|
| 554471000005108 | Danish module | DK |
| 32506021000036107 | SNOMED Clinical Terms Australian extension | AU |
| 999000011000000103 | SNOMED CT United Kingdom clinical extension module | UK |

2.1 The type of content by hierarchy

Each Top level hierarchy reviewed below for extension content.

Duplicates were looked for through comparing terms across extensions within given hierarchy. For example, "Look for duplicate terms within the procedure hierarchy". Duplicates within a module were also ignored.

Analysis was done on the complete aggregate of extensions plus the (International Core). The presence of duplication may indicate:

1. Extension concepts also in the core, either before or after.
 - Those where the concept appears in the International release after it's creation in an extension represent a maintenance burden for NRC's in the absence of a promotion process.
2. At least two countries producing similar, if not same, content. Which would suggest it's not necessarily country specific content.

Initial analysis is agnostic of description types, however analysis was further performed on just FSNs to increase likelihood of duplicate detection.

A major limitation in the approach used is that translations will (almost) be inherently unique, so comparison is dependent on English terms.

It was discovered mid analysis that a setting within the analysis database, may have caused incorrect character renderings however, this is not expected to have consequence on this analysis.

```

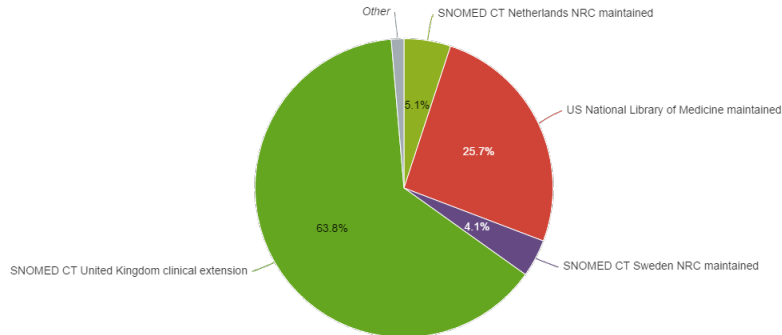
SET @Hierarchy = 404684003;

select term, count(distinct moduleId) from X_Descriptions
where conceptId in (select distinct id from X_Concepts where
active)
and moduleId != 900062011000036108 -- exclude AMT module
-- and moduleId not in(900000000000207008,90000000000012004) --
exclude international
and typeId = 900000000000003001
and conceptId in (select sourceId from X_TransitiveClosure where
destinationId = @Hierarchy)
and active = 1
group by term
having count(distinct moduleId) > 1;

-- candidates for consideration.
select * from X_Descriptions
-- active descriptions for active concepts
where active and conceptId in (select distinct id from X_Concepts
where active)
-- target heirarchy
and conceptId in (select sourceId from X_TransitiveClosure where
destinationId = @Hierarchy)
and term in (select distinct term from X_Descriptions
where conceptId in (select distinct id from
X_Concepts where active)
and moduleId != 900062011000036108 -- exclude
AMT module
-- and moduleId not in
(900000000000207008,90000000000012004)
and conceptId in (select sourceId from
X_TransitiveClosure where destinationId = @Hierarchy)
and active = 1
group by term
having count(distinct moduleId) > 1);

```

2.1.1 Clinical finding



Potential Concept Duplication

26 FSNs duplicated across the content, which are almost certainly candidates for promotion. The affected concepts are in the following extensions.

- US National Library of Medicine maintained module
- SNOMED CT United Kingdom clinical extension module
- SNOMED CT Netherlands NRC maintained module
- SNOMED Clinical Terms Australian extension
- SNOMED CT Sweden NRC maintained module
- Danish module

All but the Danish module have some overlap with each other, as well as the international release.

These are the identified FSNs.

| | |
|---|---|
| Adverse reaction to rotavirus vaccine (disorder) | Atypical atrial flutter (disorder) |
| Chronic epiglottitis (disorder) | Excessive bioactive substance intake (finding) |
| Excessive enteral nutrition infusion (finding) | Excessive growth rate (finding) |
| Excessive parenteral nutrition infusion (finding) | Family history of lactose intolerance (situation) |

Results of Analysis of SNOMED CT Extensions

| | |
|--|---|
| Inadequate bioactive substance intake (finding) | Inadequate enteral nutrition infusion (finding) |
| Inadequate parenteral nutrition infusion (finding) | Mild dementia (disorder) |
| Patient identity verified (finding) | Physical disability (finding) |
| Predicted excessive energy intake (finding) | Predicted inadequate energy intake (finding) |
| Severe dementia (disorder) | Subendocardial myocardial infarction (disorder) |
| Suspected cerebrovascular accident (situation) | Suspected sepsis (situation) |
| Takotsubo cardiomyopathy (disorder) | Thrombosis of internal jugular vein (disorder) |
| Typical atrial flutter (disorder) | Unintentional weight gain (finding) |
| Ureterostomy present (finding) | Viral meningoencephalitis (disorder) |

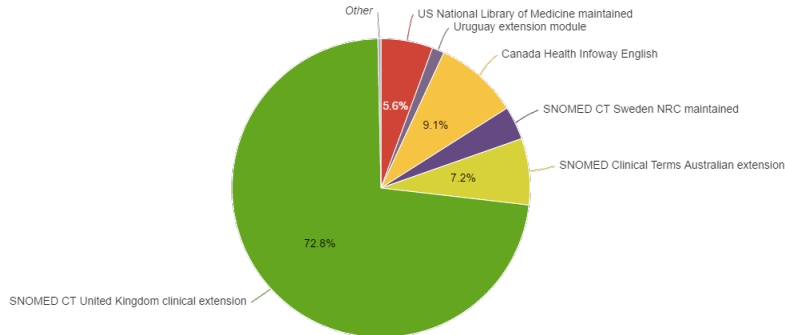
There are 6400 synonyms that are not unique across this set. There appear to be a number of reasons for this, though most seem to relate to translations.

For example:

- 371093006| Urosepsis (disorder) | has descriptions in, the extensions from three countries, that are the same as the 'en' description.
- 27830001| Brachial radiculitis (disorder) | has translations in two extensions that are different to the 'en', but differ from each other by the case of the first character.
- 75049004| Jeune thoracic dystrophy (disorder) | has translations in two extensions that appear identical.

These may have different character encoding or punctuation conventions, or written languages are genuinely similar (Danish and Swedish). A binary (eliminating case differences) compare halved the number of duplicate terms identified. It's unclear (to the author) what the standards and rules are concerning translations - are they complete (all concepts), some (only concepts of interest), as necessary (where word is different).

2.1.2 Procedure



Potential Concept Duplication

16 FSNs duplicated across the content, which are almost certainly candidates for promotion. The affected concepts are in the following extensions.

- SNOMED CT Netherlands NRC maintained module
- US National Library of Medicine maintained module
- SNOMED Clinical Terms Australian extension
- SNOMED CT United Kingdom clinical extension module
- Canada Health Infoway English module
- SNOMED CT Sweden NRC maintained module

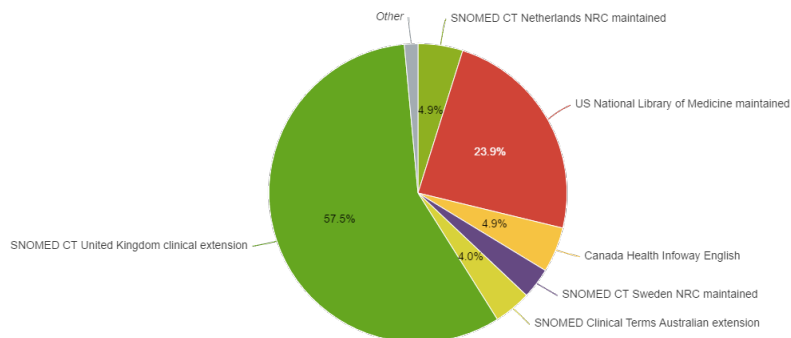
These are the identified FSNs.

| |
|---|
| Admission to nursing home (procedure) |
| Alpha-1 microglobulin measurement (procedure) |
| Autopsy planned (situation) |
| Computed tomography of head, neck, abdomen and pelvis with contrast (procedure) |
| Decompressive craniectomy (procedure) |
| electrochemotherapy (procedure) |

Results of Analysis of SNOMED CT Extensions

| |
|---|
| Flexible ureteroscopy (procedure) |
| Injection of platelet-rich plasma (procedure) |
| Laparoscopic extended right hemicolectomy (procedure) |
| Laparoscopic gastrectomy (procedure) |
| Laparoscopic radical prostatectomy (procedure) |
| Open reduction of fracture of ankle with internal fixation (procedure) |
| Open repair of strangulated incisional hernia (procedure) |
| Open repair of strangulated incisional hernia with prosthesis (procedure) |
| Rigid ureteroscopy (procedure) |
| Smoking assessment (procedure) |

2.1.3 Special concept



Potential Concept Duplication

15 FSNs duplicated across the content, which are almost certainly candidates for promotion. The affected concepts are in the following extensions.

- SNOMED CT Netherlands NRC maintained module

Results of Analysis of SNOMED CT Extensions

- US National Library of Medicine maintained module
- SNOMED CT United Kingdom clinical extension module
- Canada Health Infoway English module
- SNOMED CT core module
- Danish module
- SNOMED Clinical Terms Australian extension

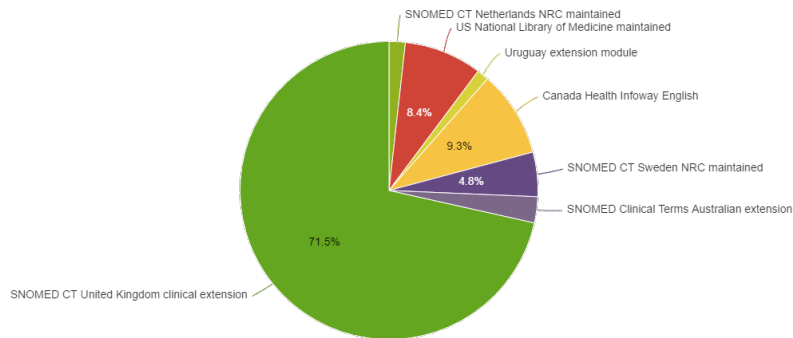
These are the identified FSNs.

| |
|---|
| 5-Hydroxyhistamine (substance) |
| Atypical atrial flutter (disorder) |
| Autopsy planned (situation) |
| Birch pollen (substance) |
| Decompressive craniectomy (procedure) |
| Deer dander (substance) |
| Escherichia coli serogroup Orough (organism) |
| Highlands j virus (organism) |
| No diabetic retinopathy (situation) |
| Subendocardial myocardial infarction (disorder) |
| Suspected cerebrovascular accident (situation) |
| Suspected sepsis (situation) |
| Takotsubo cardiomyopathy (disorder) |
| Thrombosis of internal jugular vein (disorder) |
| Typical atrial flutter (disorder) |

Results of Analysis of SNOMED CT Extensions

The mix of semantic tags in this set, suggest a possible issue with the transitive queries and history of the "aggregate release". Further investigation is required.

2.1.4 Situation with explicit context



Potential Concept Duplication

8 FSNs duplicated across the content, which are almost certainly candidates for promotion. The affected concepts are in the following extensions.

- SNOMED CT United Kingdom clinical extension module
- SNOMED CT Netherlands NRC maintained module
- SNOMED CT Sweden NRC maintained module
- US National Library of Medicine maintained module

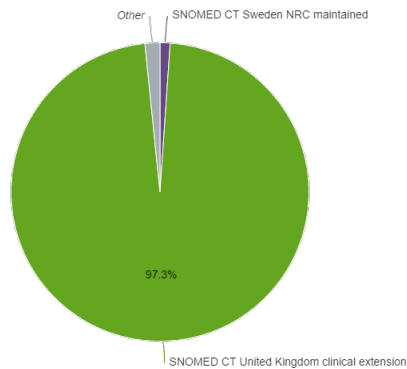
These are the identified FSNs.

| |
|---|
| Autopsy planned (situation) |
| Family history of lactose intolerance (situation) |
| History of acute coronary syndrome (situation) |
| History of amaurosis fugax (situation) |
| History of supraventricular tachycardia (situation) |
| No diabetic retinopathy (situation) |

| |
|--|
| Suspected cerebrovascular accident (situation) |
|--|

| |
|------------------------------|
| Suspected sepsis (situation) |
|------------------------------|

2.1.5 Observable entity



Potential Concept Duplication

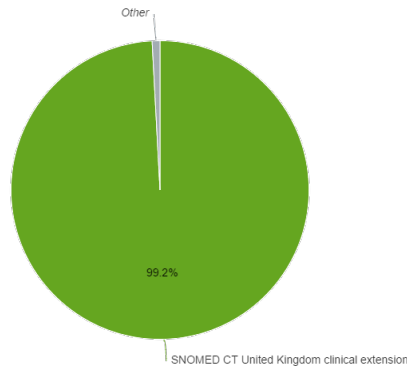
There are no FSNs duplicated across the content, which are almost certainly candidates for promotion.

There are 476 duplicate synonyms across this set. The affected concepts are in the following extensions.

- SNOMED CT Netherlands NRC maintained module
- SNOMED CT core module
- Danish module
- SNOMED CT Sweden NRC maintained module
- SNOMED CT United Kingdom clinical extension module

Results of Analysis of SNOMED CT Extensions

2.1.6 Event



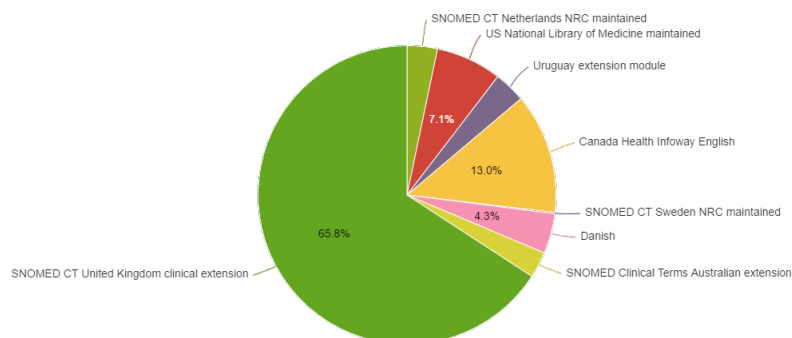
Potential Concept Duplication

No FSNs are duplicated across the content.

17 synonyms are duplicated, the affected concepts are in the following extensions.

- Danish module
- SNOMED CT core module
- SNOMED CT Sweden NRC maintained module
- SNOMED CT Netherlands NRC maintained module

2.1.7 Qualifier value



Potential Concept Duplication

28 FSNs duplicated across the content, which are almost certainly candidates for promotion. The affected concepts are in the following extensions.

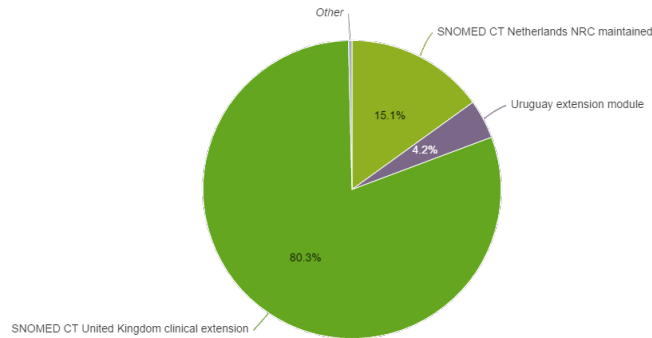
- SNOMED CT United Kingdom clinical extension module
- Canada Health Infoway English module
- US National Library of Medicine maintained module

These are the identified FSNs.

| | |
|--|---|
| Adolescent medicine service (qualifier value) | Adolescent psychiatry service (qualifier value) |
| Audiometry (qualifier value) | Cardiology service (qualifier value) |
| Chiropody (qualifier value) | Clinical allergy service (qualifier value) |
| Clinical genetics service (qualifier value) | Clinical immunology service (qualifier value) |
| Clinical neurophysiology service (qualifier value) | Dental hygiene service (qualifier value) |
| Dentistry service (qualifier value) | Dialysis service (qualifier value) |
| Endodontic service (qualifier value) | Genetics (qualifier value) |
| Health belief model (qualifier value) | Infectious diseases service (qualifier value) |
| Intensive care medicine (qualifier value) | Neonatal intensive care service (qualifier value) |
| Nephrology service (qualifier value) | Neurology service (qualifier value) |
| Physiotherapy service (qualifier value) | Prosthetics (qualifier value) |
| Prosthetics service (qualifier value) | Prosthodontic service (qualifier value) |
| Psychology (qualifier value) | Respiratory medicine service (qualifier value) |

| | |
|--|-----------------------------------|
| Respite care service (qualifier value) | Sports medicine (qualifier value) |
|--|-----------------------------------|

2.1.8 Record artifact



Potential Concept Duplication

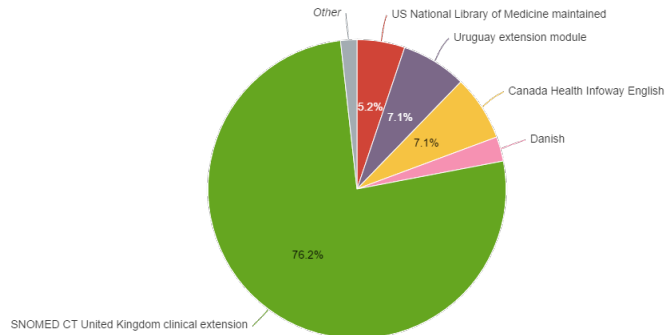
8 FSNs duplicated across the content, which are almost certainly candidates for promotion. The affected concepts are in the following extensions.

- SNOMED CT Netherlands NRC maintained module
- SNOMED CT United Kingdom clinical extension module

These are the identified FSNs.

| |
|--|
| Discharge letter (record artifact) |
| Do not attempt cardiopulmonary resuscitation order (record artifact) |
| Growth chart (record artifact) |
| Letter (record artifact) |
| Living will and advance directive record (record artifact) |
| Medical photograph (record artifact) |
| Referral letter (record artifact) |
| Weight chart (record artifact) |

2.1.9 Social context



Potential Concept Duplication

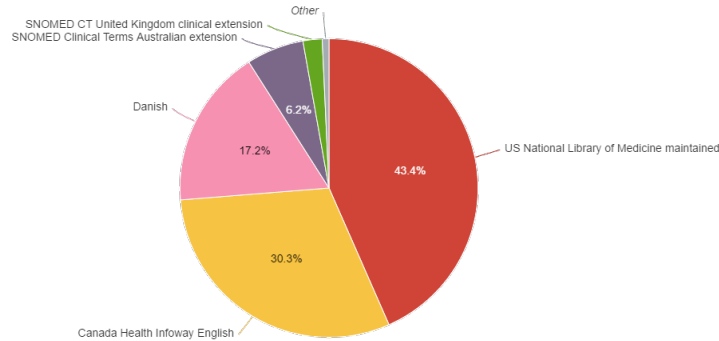
5 FSNs duplicated across the content, which are almost certainly candidates for promotion. The affected concepts are in the following extensions.

- US National Library of Medicine maintained module
- SNOMED Clinical Terms Australian extension
- Canada Health Infoway English module

These are the identified FSNs.

| |
|--------------------------------|
| Massage therapist (occupation) |
| Maternal aunt (person) |
| Maternal uncle (person) |
| Paternal aunt (person) |
| Paternal uncle (person) |

2.1.10 Substance



Potential Concept Duplication

15 FSNs duplicated across the content, which are almost certainly candidates for promotion. The affected concepts are in the following extensions.

- US National Library of Medicine maintained module
- SNOMED Clinical Terms Australian extension
- Canada Health Infoway English module
- SNOMED CT core module
- SNOMED CT United Kingdom clinical extension module

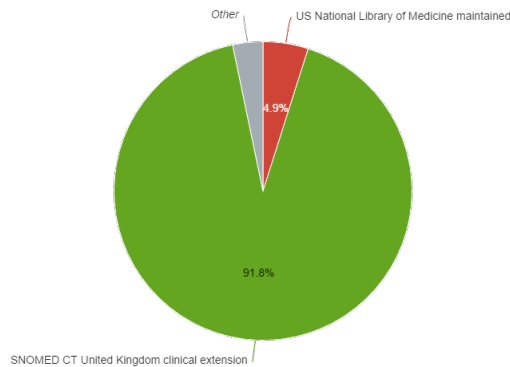
These are the identified FSNs.

| |
|--------------------------------|
| 5-Hydroxyhistamine (substance) |
| Arugula (substance) |
| Birch pollen (substance) |
| Blue cheese (substance) |
| Deer dander (substance) |
| Flounder (substance) |
| Ham (substance) |

Results of Analysis of SNOMED CT Extensions

| |
|--|
| Hickory nut (substance) |
| Honeydew melon (substance) |
| Jalapeno pepper (substance) |
| Pneumococcal conjugate vaccine (product) |
| Red onion (substance) |
| Snail - dietary (substance) |
| Tree nut (substance) |
| White pepper (substance) |

2.1.11 Body structure



Potential Concept Duplication

No FSNs duplicated across the content, which are almost certainly candidates for promotion.

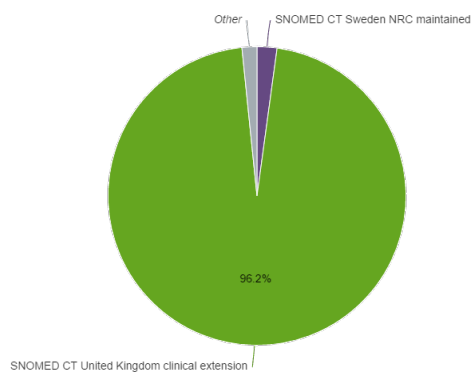
1,888 synonyms are duplicated across the content, the affected concepts are in the following extensions.

- Danish module
- SNOMED CT Sweden NRC maintained module
- SNOMED CT core module

Results of Analysis of SNOMED CT Extensions

- Lithuania
- SNOMED Clinical Terms Australian extension
- US National Library of Medicine maintained module
- SNOMED CT United Kingdom clinical extension module

2.1.12 Staging and scales

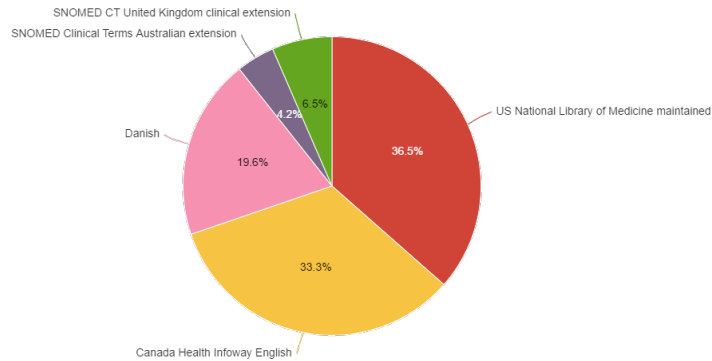


Potential Concept Duplication

No FSNs duplicated across the content, which are almost certainly candidates for promotion. 464 synonyms are duplicated across the extensions. The affected concepts are in the following extensions.

- Danish module
- SNOMED CT core module
- SNOMED CT Sweden NRC maintained module
- SNOMED CT United Kingdom clinical extension module

2.1.13 Pharmaceutical / biologic product



Potential Concept Duplication

12 FSNs duplicated across the content, which are almost certainly candidates for promotion. The affected concepts are in the following extensions.

- Canada Health Infoway English module
- US National Library of Medicine maintained module
- SNOMED Clinical Terms Australian extension

These are the identified FSNs.

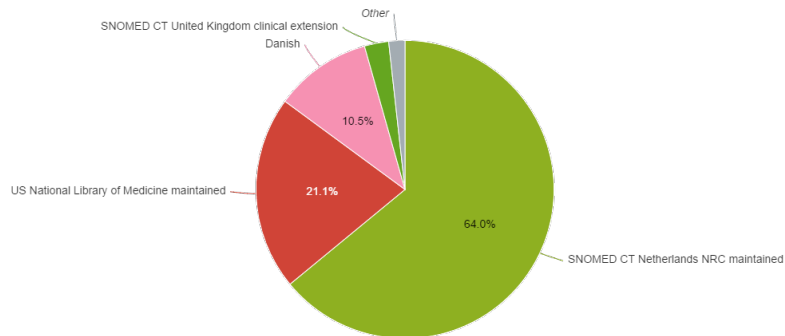
| |
|--|
| Arugula (substance) |
| Blue cheese (substance) |
| Flounder (substance) |
| Ham (substance) |
| Hickory nut (substance) |
| Honeydew melon (substance) |
| Jalapeno pepper (substance) |
| Pneumococcal conjugate vaccine (product) |

Results of Analysis of SNOMED CT Extensions

| |
|-----------------------------|
| Red onion (substance) |
| Snail - dietary (substance) |
| Tree nut (substance) |
| White pepper (substance) |

There is obviously an issue with the semantic tag and transitive queries. This may be a problem with the analysis or content.

2.1.14 Organism



Potential Concept Duplication

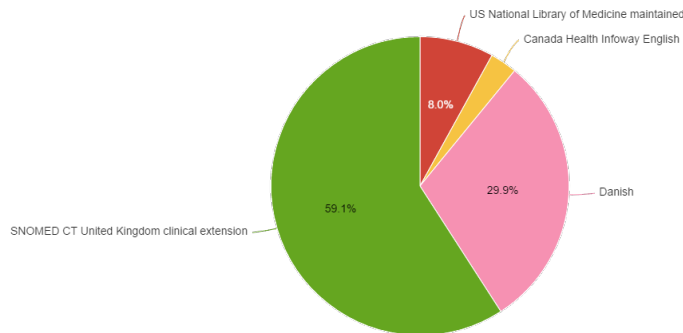
2 FSNs duplicated across the content, which are almost certainly candidates for promotion. The affected concepts are in the following extensions.

- US National Library of Medicine maintained module
- SNOMED CT core module
- Canada Health Infoway English module

These are the identified FSNs.

| |
|--|
| Escherichia coli serogroup Orough (organism) |
| Highlands j virus (organism) |

2.1.15 Environment or geographical location

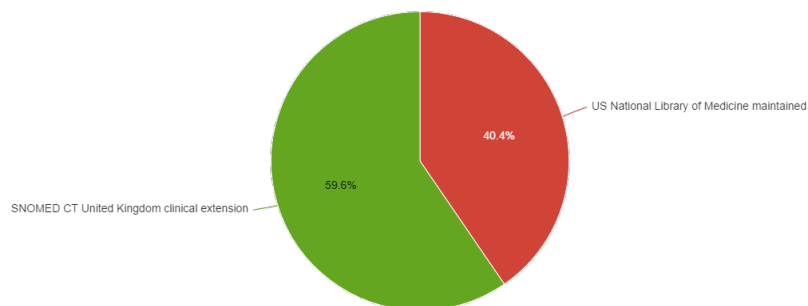


Potential Concept Duplication

No FSNs duplicated across the content, which are almost certainly candidates for promotion. 353 synonyms are duplicated across the extensions, The affected concepts are in the following extensions.

- Danish module
- SNOMED CT Sweden NRC maintained module
- SNOMED CT core module
- Lithuania

2.1.16 Specimen

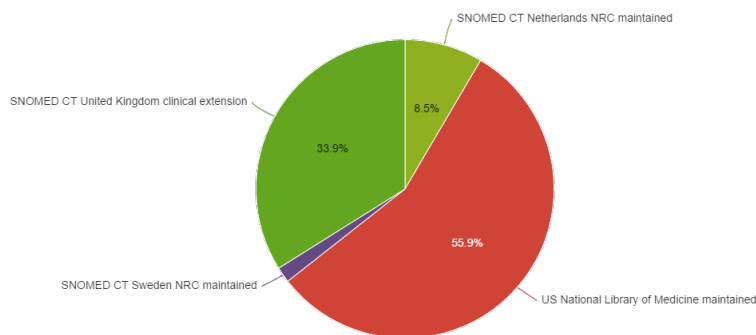


Potential Concept Duplication

No FSNs duplicated across the content, which are almost certainly candidates for promotion. Nine synonyms are duplicated. The affected concepts are in the following extensions.

- Danish module
- SNOMED CT Sweden NRC maintained module
- SNOMED CT core module
- US National Library of Medicine maintained module
- SNOMED CT United Kingdom clinical extension module

2.1.17 Physical object



Potential Concept Duplication

No FSNs duplicated across the content, which are almost certainly candidates for promotion. 399 synonyms are duplicated across the extensions. The affected concepts are in the following extensions.

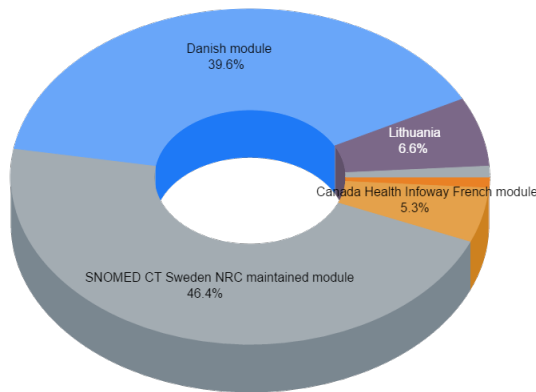
- Danish module
- SNOMED CT Sweden NRC maintained module
- SNOMED CT core module

2.1.18 Physical force

Single Concept : *U-V radiation in diagnosis NOS (physical force)*

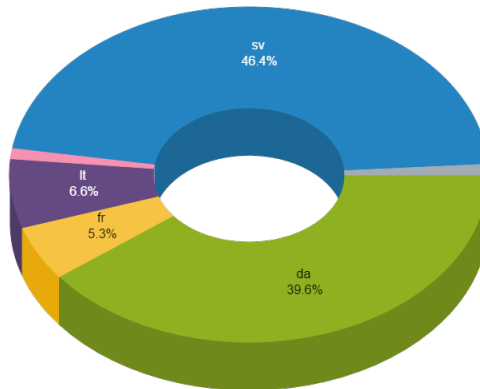
3 Extension Descriptions

Most analysis performed as part of identifying duplicates within concepts. However, below is a summary of the translations - (extension descriptions for core concepts).



| NRC | Translated Concepts |
|--|---------------------|
| SNOMED CT Netherlands NRC maintained module | 6759 |
| US National Library of Medicine maintained module | 169 |
| Uruguay extension module | 227 |
| Canada Health Infoway French module | 34063 |
| Canada Health Infoway English module | 533 |
| SNOMED CT Sweden NRC maintained module | 298331 |
| Danish module | 254752 |
| SNOMED Clinical Terms Australian extension | 1929 |
| Lithuania | 42668 |
| SNOMED CT United Kingdom clinical extension module | 3799 |

Results of Analysis of SNOMED CT Extensions



| Language | Concepts |
|----------|----------|
| da | 254752 |
| en | 6327 |
| es | 227 |
| fr | 34063 |
| it | 42668 |
| nl | 6756 |
| sv | 298331 |

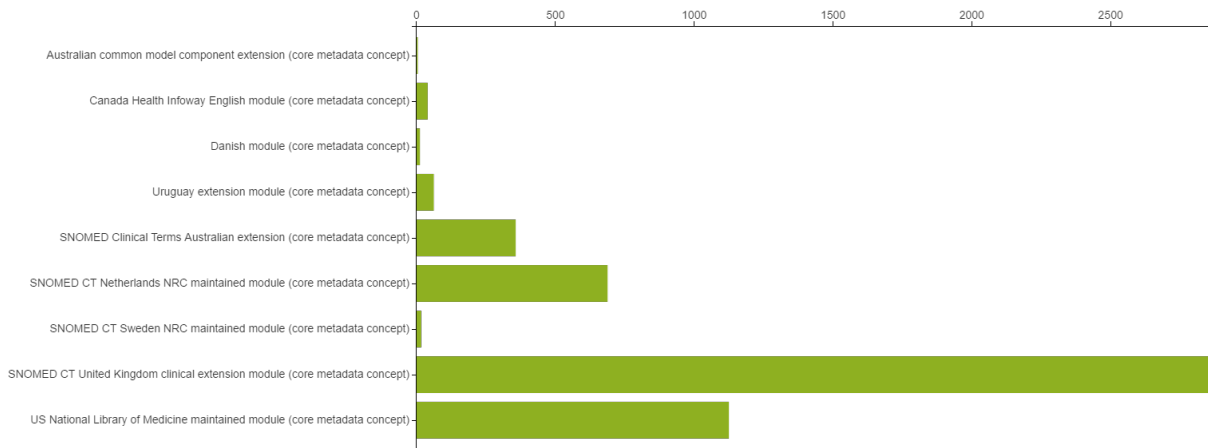
3.1 Extension Changes to Core Descriptions

178 International descriptions have some modification in an extension. The associated modules are:

- Australian common model component extension
- SNOMED Clinical Terms Australian extension
- US National Library of Medicine maintained module

4 Relationship Extensions

5,136 core concepts have been changes within an extension. Some of these look like promotions, however the majority do not appear to be.



| Extension | Inferred | Stated |
|--|----------|--------|
| Australian common model component extension (core metadata concept) | 4 | 57 |
| Canada Health Infoway English module (core metadata concept) | 40 | 0 |
| Danish module (core metadata concept) | 12 | 6 |
| Uruguay extension module (core metadata concept) | 62 | 23 |
| SNOMED Clinical Terms Australian extension (core metadata concept) | 357 | 98 |
| SNOMED CT Netherlands NRC maintained module (core metadata concept) | 688 | 0 |
| SNOMED CT Sweden NRC maintained module (core metadata concept) | 17 | 0 |
| SNOMED CT United Kingdom clinical extension module (core metadata concept) | 2862 | 2848 |
| US National Library of Medicine maintained module (core metadata concept) | 1125 | 401 |

Note: Some of the numbers comparing stated and inferred look odd, this is likely a result of the crude aggregation of extensions and some of the extension content already having been promoted to core.

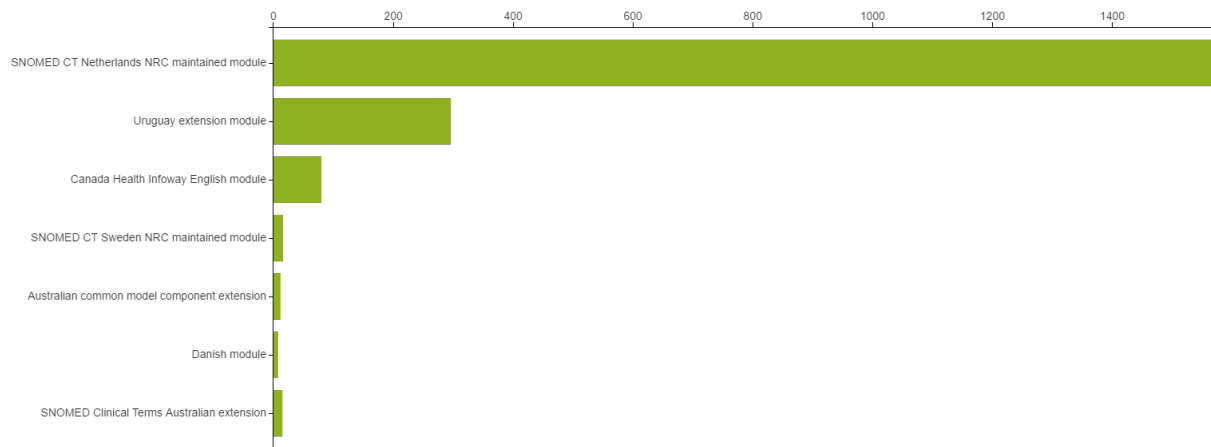
4.1 Core relationships modified within an Extension

1,997 core relationships were modified by an extension, affecting 384 concepts

A single concept, 425630003| *Acute irritant contact dermatitis (disorder)* | was modified by two NRCs.

Both inactivated all the relationships, but one recreated them in the subsequent release.

Other changes are summarised below.

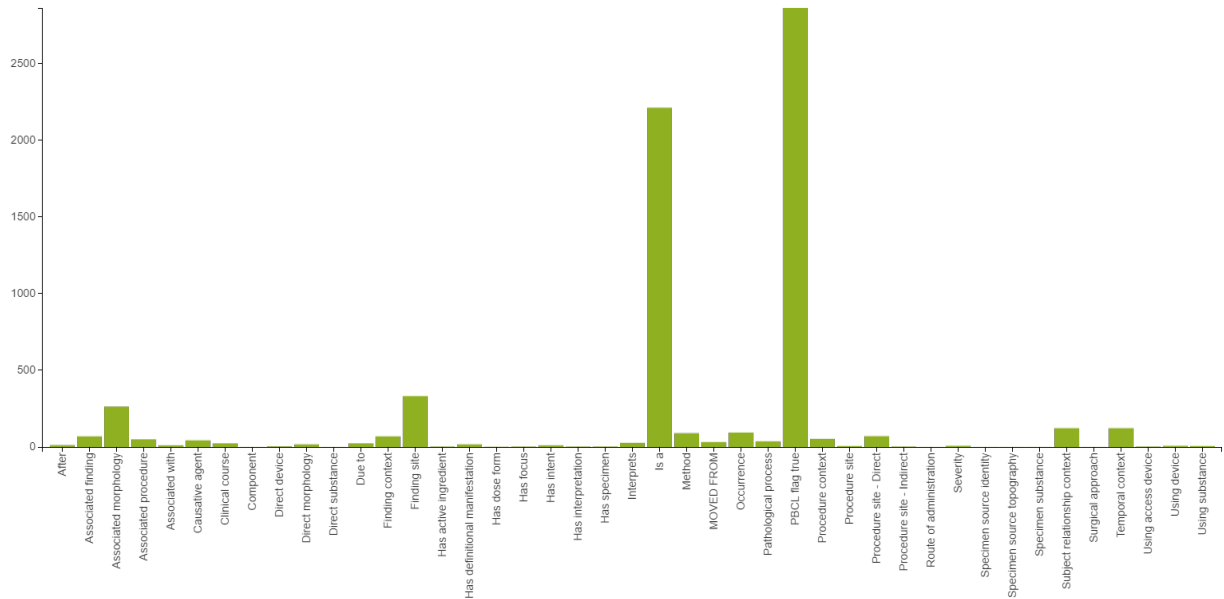


| NRC | Relationships Changed |
|---|-----------------------|
| SNOMED CT Netherlands NRC maintained module | 1570 |
| Uruguay extension module | 296 |
| Canada Health Infoway English module | 80 |
| SNOMED CT Sweden NRC maintained module | 16 |
| Australian common model component extension | 12 |
| Danish module | 8 |

| | |
|--|----|
| SNOMED Clinical Terms Australian extension | 15 |
|--|----|

4.2 Types of Relationships Modified

A large variety (43) of relationship types are involved in the edits, most are *ISA*, and some are not part of the approved concept model or are attributes specific to an extension.



| Relationship Type | Number of relationships |
|-----------------------|-------------------------|
| After | 12 |
| Associated finding | 69 |
| Associated morphology | 263 |
| Associated procedure | 50 |
| Associated with | 11 |
| Causative agent | 43 |
| Clinical course | 24 |

Results of Analysis of SNOMED CT Extensions

| | |
|--------------------------------|------|
| Component | 1 |
| Direct device | 4 |
| Direct morphology | 17 |
| Direct substance | 1 |
| Due to | 24 |
| Finding context | 69 |
| Finding site | 331 |
| Has active ingredient | 3 |
| Has definitional manifestation | 17 |
| Has dose form | 2 |
| Has focus | 3 |
| Has intent | 11 |
| Has interpretation | 3 |
| Has specimen | 3 |
| Interprets | 27 |
| Is a | 2209 |
| Method | 90 |
| MOVED FROM | 32 |
| Occurrence | 94 |
| Pathological process | 37 |

Results of Analysis of SNOMED CT Extensions

| | |
|------------------------------|------|
| PBCL flag true | 2860 |
| Procedure context | 53 |
| Procedure site | 6 |
| Procedure site - Direct | 70 |
| Procedure site - Indirect | 3 |
| Route of administration | 1 |
| Severity | 7 |
| Specimen source identity | 1 |
| Specimen source topography | 1 |
| Specimen substance | 1 |
| Subject relationship context | 122 |
| Surgical approach | 1 |
| Temporal context | 122 |
| Using access device | 3 |
| Using device | 7 |
| Using substance | 6 |

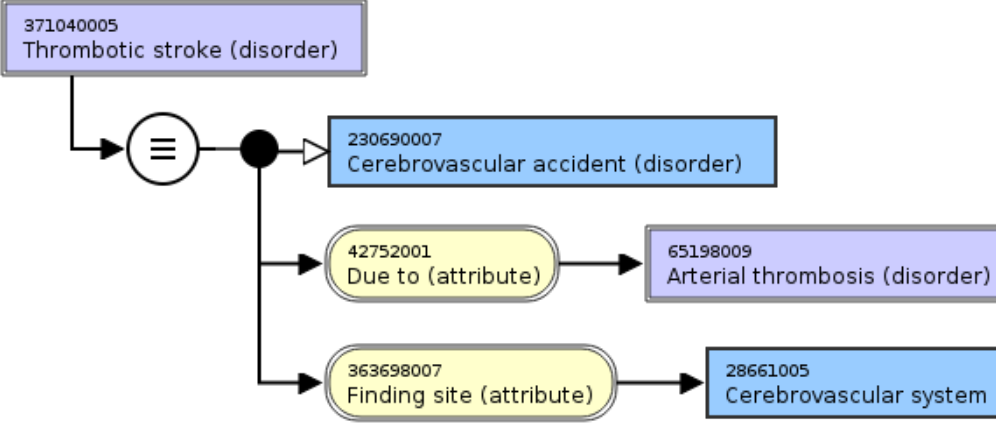
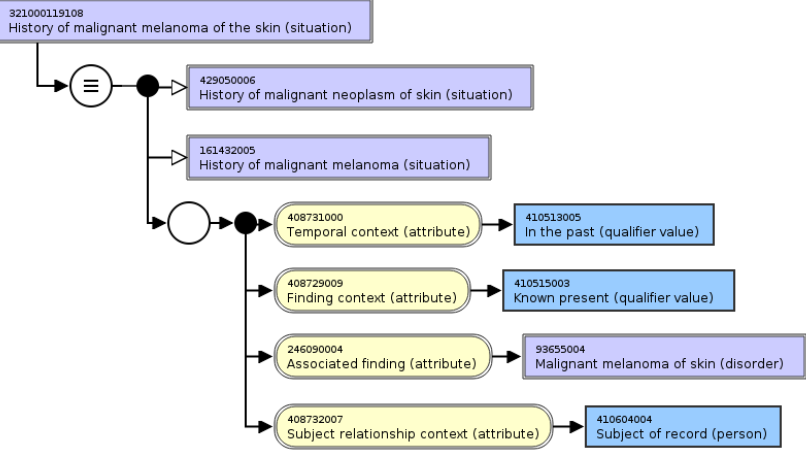
```

select count(distinct sourceId)  from X_Relationships
where moduleId not in(900000000000207008,90000000000012004,900062
011000036108) -- exclude international+AMT
and sourceId in (select id from X_Concepts where moduleId in(9000
00000000207008,90000000000012004) and active)
group by moduleId;

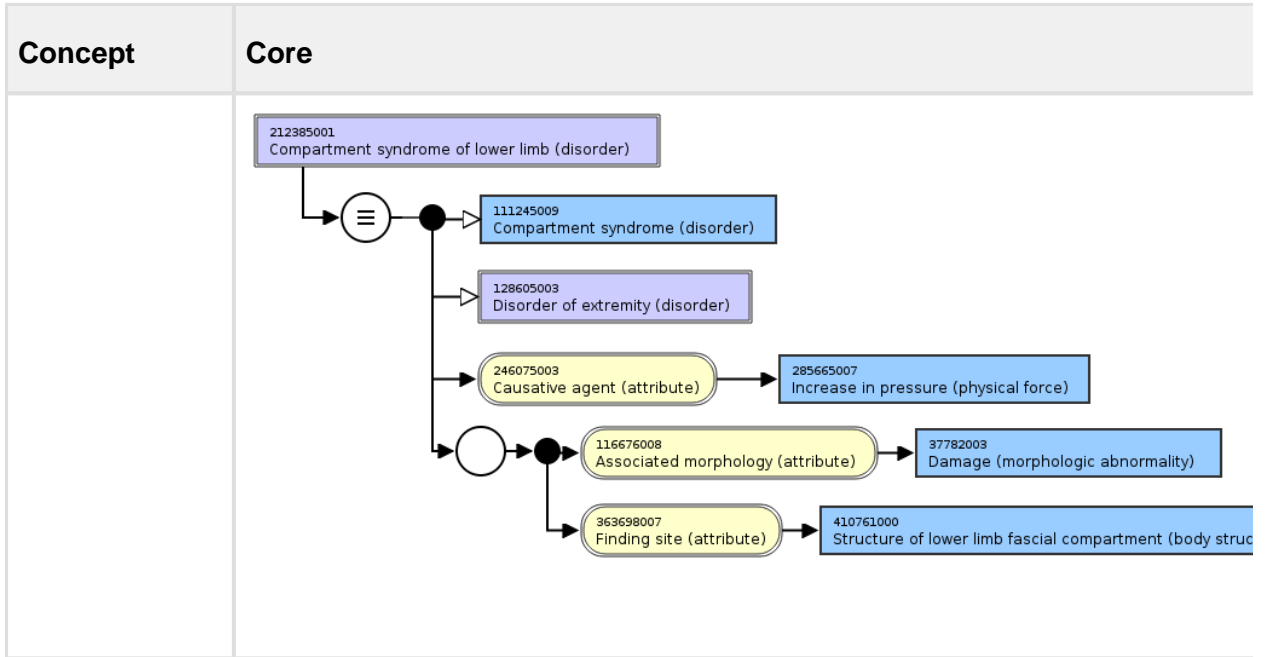
```

Results of Analysis of SNOMED CT Extensions

Comparison Examples - Published (inferred) relationships for Core concepts

| Concept | Core |
|--------------|--|
| 371040005 |  <pre> graph TD A["371040005 Thrombotic stroke (disorder)"] --> B((≡)) B --> C((●)) C --> D["230690007 Cerebrovascular accident (disorder)"] C --> E["42752001 Due to (attribute)"] C --> F["363698007 Finding site (attribute)"] E --> G["65198009 Arterial thrombosis (disorder)"] F --> H["28661005 Cerebrovascular system"] </pre> |
| 321000119108 |  <pre> graph TD A["321000119108 History of malignant melanoma of the skin (situation)"] --> B((≡)) B --> C((●)) C --> D["429050006 History of malignant neoplasm of skin (situation)"] C --> E["161432005 History of malignant melanoma (situation)"] C --> F((○)) F --> G["408731000 Temporal context (attribute)"] F --> H["408729009 Finding context (attribute)"] F --> I["246090004 Associated finding (attribute)"] F --> J["408732007 Subject relationship context (attribute)"] G --> K["410513005 In the past (qualifier value)"] H --> L["410515003 Known present (qualifier value)"] I --> M["93655004 Malignant melanoma of skin (disorder)"] J --> N["410604004 Subject of record (person)"] </pre> |
| 212385001 | |

Results of Analysis of SNOMED CT Extensions



5 Additional Observations

The following details are only exemplars of the observations made, and by no means comprehensive.

6 Extensions vs Editions

Of the 9 releases looked at:

- Four published [Editions](#)
- Three publish [Extensions](#)
- One publishes three separate extensions.
- One publishes an extension, "bundled" with the International Edition.

6.1 File naming

The file naming conventions, do not appear to be consistent across the extensions.

- sct2_Concept_Snapshot_AU1000036_20161231.txt
- sct2_Concept_Snapshot_en-CanadianExtension_20161031.txt
- sct2_Concept_Snapshot_DK1000005_20161130.txt
- sct2_Concept_Snapshot_LT1000092_20151107.txt
- sct2_Concept_Snapshot_NL_20160930.txt
- sct2_Concept_Snapshot_SE1000052_20161130.txt
- sct2_Concept_Snapshot_GB1000000_20161001.txt
- sct2_Concept_Snapshot_US1000124_20160901.txt
- sct2_Concept_Snapshot_es-UruguayExtension_20161215.txt
- sct2_Concept_Snapshot_INT_20160731.txt

6.2 Directory structure

Some variation was noticed in the the directory structure within the published zip files.

Below are the paths the the snapshot concepts file in each release.

- \SnomedCT_Release_AU1000036_20161231\RF2Release\Snapshot\Terminology
- \SnomedCT_Canadian_EnglishExtension_Release_20161031\Snapshot\Terminology
- \SnomedCT_ManagedServiceDK_Production_DK1000005_20161130\Snapshot\Terminol
- \SnomedCT_RF2Release_LT1000092_20151107\Snapshot\Terminology
- \SnomedCT_Netherlands_EditionRelease_20160930\Snapshot\Terminology

Results of Analysis of SNOMED CT Extensions

- \SnomedCT_SE_Production_20161130T170000\Snapshot\Terminology
- \SnomedCT_RF2Release_GB1000000_20161001\Snapshot\Terminology
- \SnomedCT_RF2Release_US1000124_20160901\Snapshot\Terminology
- \SnomedCT_Uruguay_Extension_Release_20161215\Snapshot\Terminology

6.3 Specific file inclusions

The international release includes 6 files - Concept, Description, Relationship, StatedRelationship, Identifier and TextDefinition files - within the "Terminology Folder"
The files are not consistently present in extensions.

| | Concept | Description | Relationship | StatedRelationship | Identifier | Text |
|-------------|---------|-------------|--------------|--------------------|------------|------|
| Australia | 1 | 1 | 1 | | 1 | |
| Candana* | 1 | 1 | 1 | 1 | | 1 |
| Denmark** | 1 | 2 | 1 | 1 | 1 | 2 |
| Lithuania | 1 | 1 | 1 | | | |
| Netherlands | 1 | 1 | 1 | | | 1 |
| Sweden** | 1 | 2 | 1 | 1 | 1 | |
| UK | 1 | 1 | 1 | 1 | | |
| USA | 1 | 1 | 1 | 1 | 1 | 1 |
| Uruguay | 1 | 1 | 1 | 1 | | 1 |

* Canada include a French and English bundle.

** Sweden and Denmark include both an English and Native language Description file.

Denmark and USA are the countries to include all 6 files.

Further variations are present amongst within the *reset* subdirectory.

6.4 Miscellaneous QA issues

The description file for one extension for found to be missing the language code for 268 entries.
(The country was notified and have rectified).

7 Conclusion

- The analysis described above reveals a wealth of information. There is evidence of **duplication of content in almost every hierarchy**, the extent of which is likely to be **much greater** given the primitive analysis techniques used.
- There is value to the whole SNOMED CT community to **introduce a process for content to be promoted through to core**. This process should honour the identifiers issued by an extension builder, so as to **minimise maintenance burden of the originating extension**, and recognise their effort. The **current process is prohibitive to promoting content**, and consequently content is duplicated across extensions, the potential maintenance debt grows.
- Almost **all NRCs are actively modifying core content**. Which proves the importance of clarifying the issues raised in the Discussion Paper - Allowance of Extensions to Modify Core Content ([SNOMED International Response](#)). It seems Members have taken a different interpretation of the license to that held by the governing body, and this discrepancy has never been recognised.
- Variations in the artefacts published by NRCs exist. No comment is made about compliance with Technical Specifications, but such variances may impact portability of software that consumes SNOMED CT. **Some sort of certification/verification process asserting a minimum conformance criteria would prove valuable.**
- All the issues described here may just as likely apply to other (affiliate) extensions, however will remain unknown without systematic investigation.
- The issues described are real and Members are currently struggling to deal with. Prolonging their resolution introduces a cost to all.

8 Endnotes

¹The author recalls a requirement that a US spelling 'en' FSN should be created for all concepts, but unable to identify this in current specification. Is this still a requirement?