Analyzing Submitted Requests from the US SNOMED CT Content Request System (USCRS)

Project Lead

Erin Foster, NLM Associate Fellow

Project Sponsors

Suzy Roy, Librarian
Steve Emrick, Head Terminology QA & User Services Unit

MEDLARS Management Section (MMS)

25 August 2015

Table of Contents

Abstract

Objective

The primary objective of this project is to analyze requests submitted to the United States SNOMED CT Content Request System (USCRS) to gain a better understanding of these requests. This exploratory data analysis provides insight into the types of requests submitted, submitter characteristics, and the timeline for processing requests in USCRS. This project also informs future development of the USCRS system.

Methods

The research used quantitative analysis to analyze requests submitted to USCRS. Prior to retrieving data from USCRS, research questions were developed to guide retrieval of data from the USCRS database. Data pulled from the USCRS database was then analyzed to answer these research questions and determine areas for improvement of USCRS's workflow and information capture.

Results

The primary results from the analysis of USCRS request data included: identification of submitter characteristics, examination of rejected requests and associated documentation, as well as determination of the average time for a subset of requests submitted to USCRS to complete the request process.

Discussion

Challenges encountered during this project primarily originated from limitations of the USCRS system in capturing information about requests submitted to the system. As a result of these challenges, recommendations are made with an emphasis on system requirements and enhancing future efforts to capture data submitted to USCRS.

Conclusion

Moving forward, the results of this project provide NLM with better insight into the requests included in USCRS. The challenges encountered during this project inform future development of the USCRS system in terms of functionality and scope. This project offers a foundation for future research and analysis of USCRS content perhaps through a more structured data analysis. The research performed also offers a sample methodology for analysis of other terminology and vocabulary request systems developed and maintained by NLM.

Background

The United States SNOMED CT Content Request System (USCRS) is a system developed by the National Library of Medicine (NLM) to manage requests from the United States for basic changes to SNOMED CT. SNOMED CT is a global clinical terminology that covers a broad range of clinical specialties and disciplines. The International Health Terminology and Standards Development Organisation (IHTSDO) owns, maintains, and provides access to SNOMED CT. [1] IHTSDO is an international non-profit organization comprised of 27 member countries. These member countries provide free access to SNOMED CT within their countries and maintain national extensions of SNOMED CT. The United States (US) is a member country of IHTSDO and NLM maintains the US Extension to SNOMED CT, which is comprised of concepts, descriptions, relationships and their history that is pertinent to health care practice, systems, and professionals within the United States. [2] The US Edition is released twice a year and is comprised of the US Extension and the International release of SNOMED CT. [3]

USCRS was initially released in 2010 and currently contains over 18,000 requests. The system was created to facilitate the development and distribution of the US Extension of SNOMED CT as well as contribute to the International version of SNOMED CT. [4] The content of USCRS is categorized broadly as requests to "add", "change", or "retire" information in SNOMED CT. These categories are further broken down into eleven areas: *New Concept*, *New Synonym*, *Add Parent*, *Change Description*, *Change Parent*, *Change Relationship*, *New Relationship*, *Retire Concept*, *Retire Description*, *Retire Relationship*, and *Other*. Users must register for a UMLS Terminology Services (UTS) account before submitting requests to USCRS and are required to have an understanding of SNOMED CT's structure and content. When submitting to the system, users must provide justification or a practical use case for changes requested as well as link to appropriate identifiers in either SNOMED CT International or the US Extension of SNOMED CT.

Once requests are submitted to USCRS, they undergo a review process to determine whether or not they will be included in the US Edition/Extension or forwarded onto SNOMED CT International. There are 12 states for requests including: accepted, appeal rejected, approved, awaiting policy, clarification, completed, forwarded, being modeled, new, rejected, retracted, under appeal. Of these, 3 are considered "end" states – completed, rejected, and retracted. End state means that once requests are assigned one of these 3 states they are done with the review process. This project specifically focuses on the end states of "completed" and "rejected". See Appendix I – USCRS state transition diagram for an illustration of the current workflow of USCRS.

The primary motivation behind this project is to gain a better understanding of the content (i.e., requests) contained in USCRS. USCRS is an important tool used by NLM to support the national and international development, enhancement, and distribution of SNOMED CT. Better understanding of the requests submitted to this system can aid NLM in content strategy development as well as inform NLM of the value of licensing SNOMED CT for national use. NLM is also in the process of developing a new version of the USCRS system and this project is also expected to inform the development of that system.

Project Objectives

The objective of the project is to perform analysis of the requests submitted to the US SNOMED CT Content Request System (USCRS). This exploratory data analysis provides insight into submitter characteristics, rejected requests, and the lifecycle of completed requests in USCRS. Based on analyses conducted, recommendations are provided for future system development as well.

Methods

Analysis of submitted requests to USCRS involved the development of research questions to guide data retrieval from the USCRS database, data clean-up included normalization, integration, and deduplication of the raw data, and final analysis of the quantitative data using Microsoft Excel 2010.

The research questions focused on three areas – submitter characteristics, rejected requests, and determination of the lifecycle of completed requests in USCRS. The list of specific questions is included in Appendix II – Research questions.

Submitter characteristics, in this context, refer to the organizations submitting requests to USCRS. While USCRS ties each request to an individual, organizational affiliation for these individuals is usually provided and allows for a larger perspective on which companies and/or institutions engage with SNOMED CT development and use in the United States. Rejected requests were analyzed to determine how requests reach the "rejected" end state. Finally, a sample of completed requests were evaluated to determine the average time for a request to go from a "new" (i.e., submitted) to "completed" state.

Data retrieval involved querying the USCRS database through use of Standardized Query Language (SQL). The research questions guided the type of information pulled from the system; also, data collection was scoped to all requests submitted before or on March 31, 2015 in order to maintain a consistent sample. For a list of queries, please see Appendix III – SQL queries.

Following data retrieval, the raw data was organized into spreadsheets and subjected to data cleaning processes. The majority of these processes included normalization, integration, and deduplication of the raw data retrieved. The quantitative data was then analyzed using Microsoft Excel 2010.

Results & analysis

A total of 16,106 requests were submitted prior to or on March 31, 2015. The results fell into three categories: submitter characteristics, rejected requests, and the lifecycle of completed requests.

Submitter characteristics

Of the total corpus of 16,106 requests, 67 organizations submitted 15,745 requests to USCRS. Thirteen individuals who did not supply an organizational affiliation submitted the remaining 361 requests to USCRS. The submitting organizations were distilled into 9 categories: Companies, Academic Institutions, Professional Organizations, Government Agencies, Consulting Companies, Collaboratives, Unknown, State Agencies, and Committees. Figure 1 uses color to represent the different categories and size to

indicate the number of requests per type of organization as part of the total number of requests submitted to USCRS.

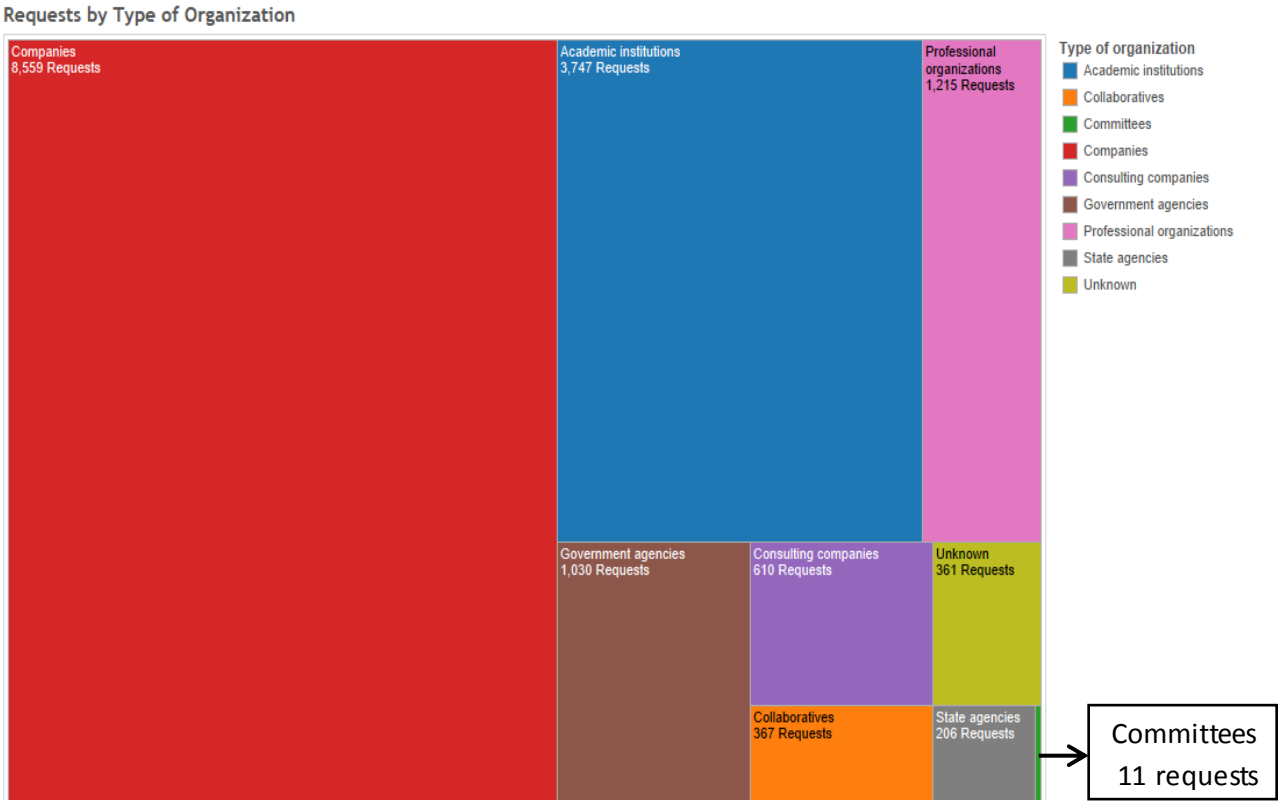


Figure 1: Requests by type of organization

For a full list of organizations with number of requests and organizations' titles per category, please see Appendix IV – Full list of submitter organizations.

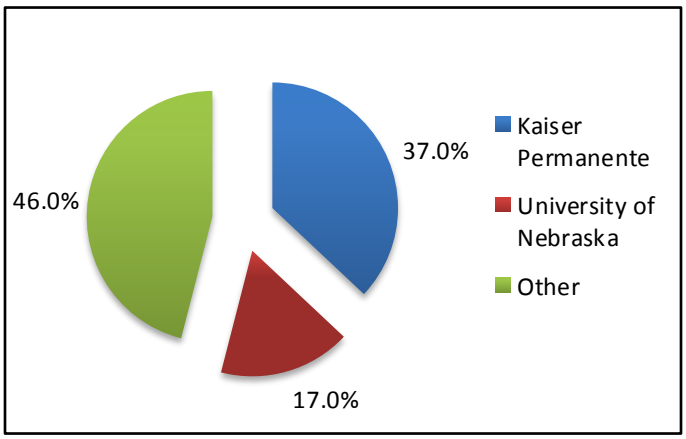


Figure 2: Submitting organizations

As Figure 1 indicates, the majority of requests (53%) were submitted by Companies (i.e., Kaiser Permanente, Intelligent Medical Objects (IMO)) with the next highest majority (23%) submitted by Academic Institutions. However, within these two categories, 54% of all requests were submitted by two organizations - Kaiser Permanente (within Companies) and University of Nebraska (within Academic Institutions) (see Figure 2). Further analysis of these requests provides insight into when these two organizations submitted the majority of their requests to USCRS.

Figure 3 is a timeline of requests submitted to USCRS for the years 2011, 2012, 2013, and 2014. These years were used because USCRS contains full calendar year data for these years. 2011 and 2014 were the highest years of submission with 7,069 requests and 5,531 requests submitted respectively. These are also the years that Kaiser Permanente and University of Nebraska made their biggest contributions. Kaiser submitted 4,251 requests in 2011 and University of Nebraska submitted 2,739 requests in 2014.

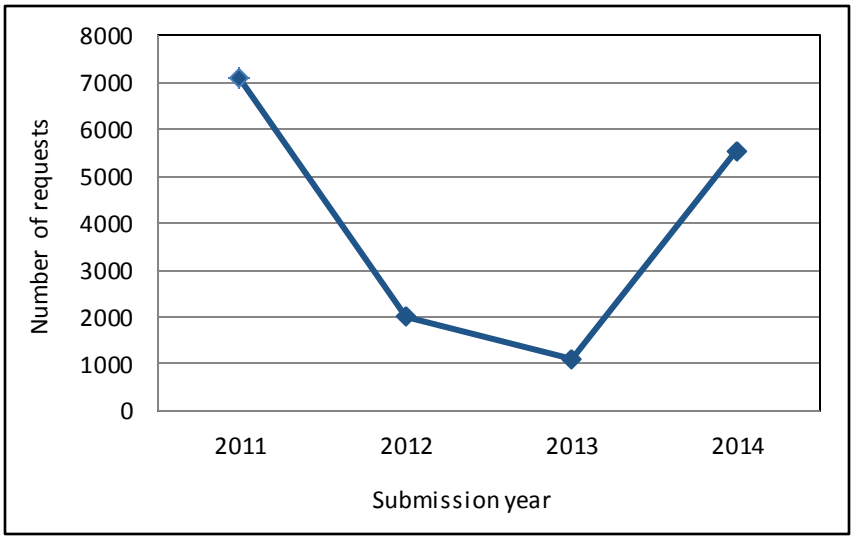


Figure 3: Requests submitted to USCRS by year

Following the request breakdown by year, request data was then used to determine the top submitters to USCRS by year (Figure 4). The small multiples chart below lists the top five submitters by year. Years 2010 and 2015 are again excluded because they are incomplete years. The colors represent organizations, which are named in the accompanying legends.

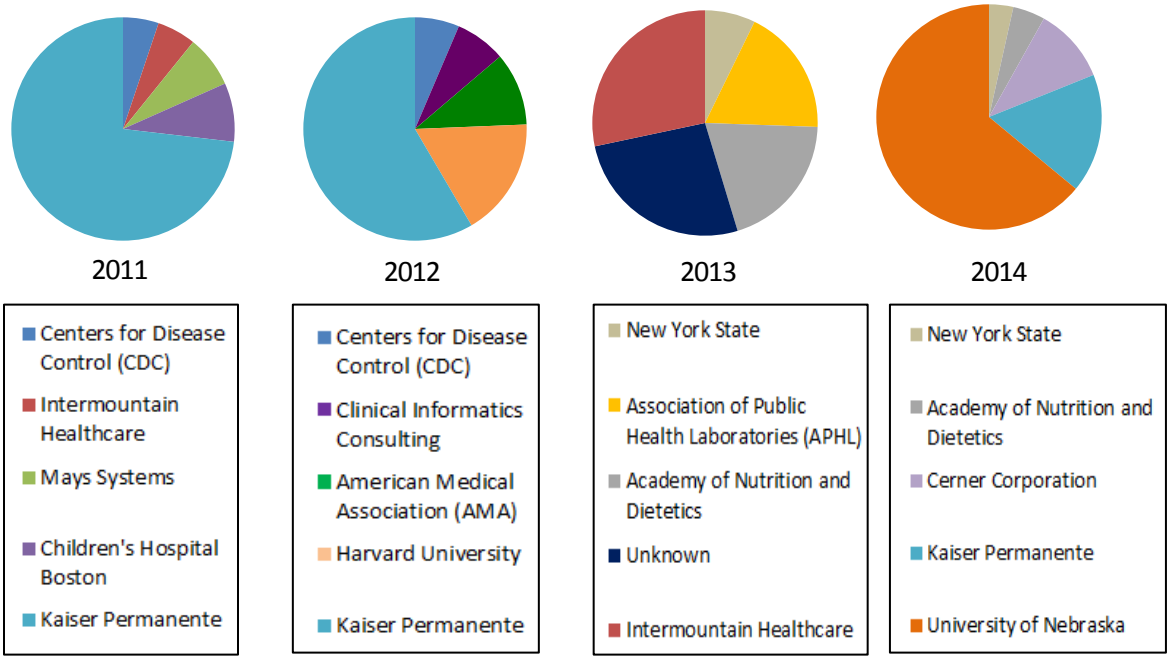


Figure 4: Top 5 submitting organizations by year

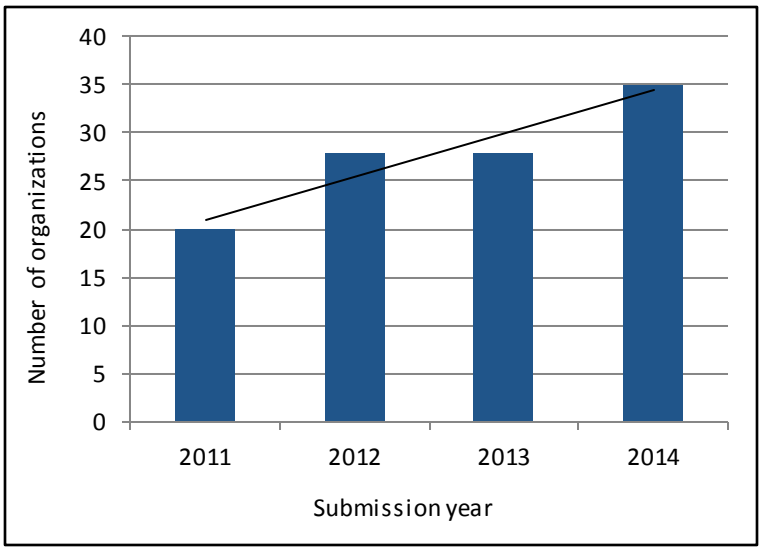


Figure 5: Organizations submitting to USCRS

Additionally, Figure 5 shows the number of organizations submitting to USCRS by year – the number has increased over time with 20 submitting organizations in 2011, 28 submitting organizations in 2012 and 2013, and finally with 35 submitting organizations in 2014. This graph also indicates an upward trend in the number of organizations submitting to USCRS.

Overall, this set of results gives a sense of the main contributors to USCRS. They demonstrate as well, particularly Figure 4, the frequency with which certain organizations submit. For example, Kaiser Permanente is a consistent contributor, landing in the top 5 submitting organizations for all years excepting 2013. Some of the fluctuation in Kaiser's submission over the years is reflective of agreements with the IHTSDO where Kaiser, following 2011, frequently submits requests for additions to SNOMED CT directly to IHTSDO rather than going through the USCRS system. Also, there are certain organizations (e.g., Harvard University) that submit large amounts of specific content for inclusion in SNOMED CT at one time, but will most likely never submit a large quantity of requests again. This also speaks to outside factors that can influence the top submitting organizations as well as the number of requests received in a given year – for example, legislation around Meaningful Use could impact the amount of requests submitted to USCRS depending on the year and the type of request submitted. Further analyses might work to identify these potential external factors and track the way they impact, if at all, submissions to USCRS. Future tracking of organizations submitting to USCRS can also open up opportunities for collaboration and partnership.

Rejected requests

This set of results is focused on requests that have been rejected as part of the USCRS request review process. There are 2,275 rejected requests available for analysis prior to or on March 31, 2015. These requests make up 14% of the total requests in USCRS (Figure 6). Very little analysis has been done of these requests prior to this project – the following graphs and charts focus on rejected requests in terms of types of requests rejected and categorization of reasons for

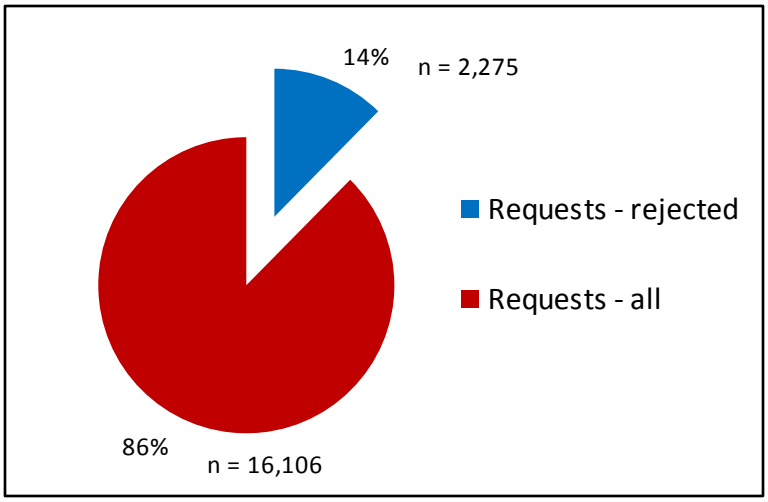


Figure 6: Rejected requests (% of total)

rejection.

As mentioned in the Background, there are 11 types of requests that can be submitted to USCRS. The graph below indicates how rejected requests break down across the different types.

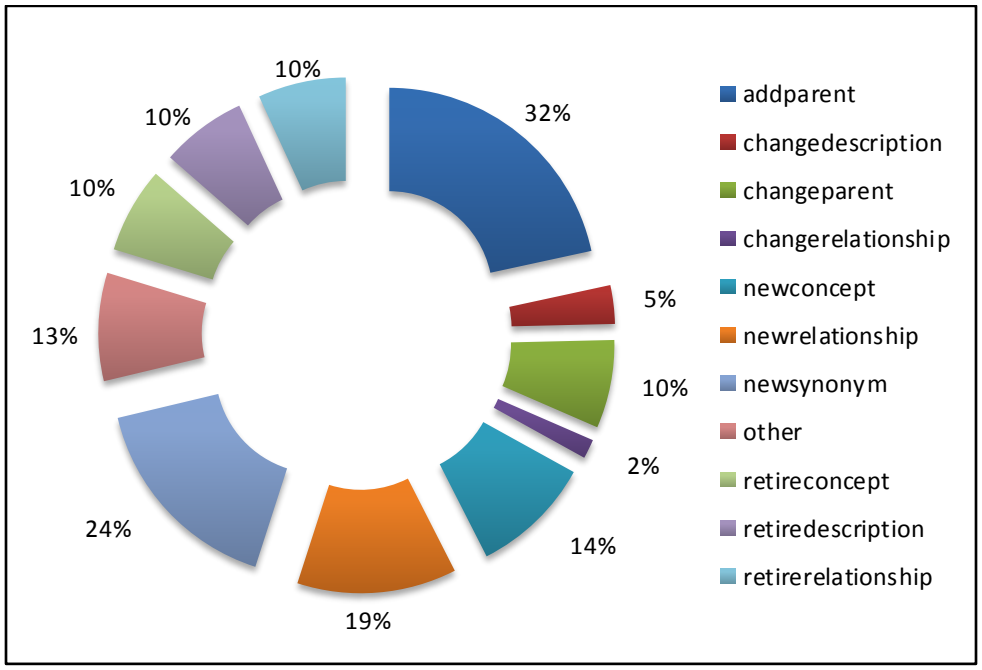


Figure 7: Types of requests rejected

An interesting aspect of Figure 7 is the relatively high percentage of certain types that are rejected particularly in context of how many are actually submitted to USCRS. *New Concept* types of requests are by far the majority of requests submitted to USCRS (93%), which means that the other 10 types each make up anywhere from .5-1% of requests submitted. Further evaluation of rejected requests may focus specifically on those types that have relatively high levels of rejection in relation to their normal levels of submission (i.e., *Add Parent*, *New Synonym*, *New Relationship*).

Another piece of the analysis of rejected requests involved evaluation of the "notes" included in request records. The notes section is a space where submitters can provide additional information about their request and where the USCRS manager (Dr. James Case) can communicate about the request (e.g., request clarification, comment on request status). If requests are rejected, these notes are the primary source of information as to why.

While notes are not a required element of USCRS records, the majority of rejected requests have them. Based on this, a sample of notes from rejected requests was selected for review and analysis. This sample was determined based off the total number of rejected requests (2,275) with an 80% confidence level and a 5% margin of error. This resulted in a sample size of 154 requests. Random numbers were generated and applied to the rejected requests to reach the final, randomized sample.[1] This sample of

[1] The randomizer generated 1 duplicate number making the final sample size 153 requests

requests was then reviewed with an eye toward discerning themes of rejection. The table below groups the final sample of 153 requests based on reason for rejection:

Reasoning	Number of requests
IHTSDO response (e.g., frequency of use)	72
Concept exists	17
Other (e.g., supernumerary structures)	15
Unclear, clarification needed	10
Refer to LOINC (e.g., observable entities)	10
Addition, separation of requests	7
Duplicate request	7
Permission not received from the American Joint Committee on Cancer (AJCC)	6
Inactive, no longer needed	4
Excessive precoordination	3
Blank	2

Analysis of these notes was limited to immediately apparent or identifiable themes/reasons – in future analyses of notes associated with rejected requests, it might be interesting to explore fundamental reasons as to why requests were rejected (e.g., lack of understanding of SNOMED CT structure, parent/child relationships, etc.). A former NLM Associate, J. Caitlin Sticco, developed an error taxonomy to code errors in requests submitted to USCRS. Further exploration of that error taxonomy and its potential for use in evaluating these notes might be warranted.

Most of these reasons are self-explanatory; however, some warrant further explanation. The majority of notes evaluated in this sample were rejected due to a response from IHTSDO. Of these, the primary reason for rejection was classified as "frequency of use = 0", which means that the concept submitted to USCRS was not evidenced to be used in the context of its submitting organization. All of these requests are encouraged to be resubmitted provided they can demonstrate use. Generally, a request will be included in the US Edition regardless of its rejection from IHTSDO; however, inability to demonstrate a frequency of use is a reason for rejection within USCRS as well. The second reason, "Concept exists", is not surprising – in fact, prior to evaluation of these notes, this reason was hypothesized to be a primary reason for rejection. The category, "Refer to LOINC", refers to those requests that were rejected because they submitted content (i.e., observable entities) that does not belong in SNOMED CT, but is more appropriate for the Logical Observation Identifiers Names and Codes (LOINC) terminology. Many of these reasons can also be contextualized within certain events – for example, those rejected with the reasoning "Permission not received from the American Joint Committee on Cancer (AJCC)" were specific to discussions between NLM and the AJCC. Finally, "Blank" refers to those requests that did not have notes attached therefore the reasons as to why these requests were rejected cannot be immediately determined.

Lifecycle of completed requests

A final set of analyses reviewed the lifecycle of completed requests within USCRS, primarily by determining how requests (in terms of time) made it from initial submission to the "completed" end

state in USCRS. In order to determine this lifecycle, a sample of completed requests was evaluated. To determine a sample, all completed, *New Concept* requests were queried for and pulled from the USCRS database. However, many of these requests did not have associated SNOMED CT Concept IDs or effective times as part of the request data, which is information needed to complete the analysis. Following this discovery, the Unified Medical Language System Terminology Services (UTS) Application Programming Interface (API) was used to run the initial data recovered from the USCRS database against the Unified Medical Language System (UMLS) in order to retrieve the information missing from these requests. After the process was complete, the data was exported and final clean-up was performed (e.g., removing multiples). Additionally, after clean-up, the days to completion were calculated (effective date minus submitted date) and some resulted in a negative number – generally because they were synonyms that had been erroneously submitted as *New Concept* requests – so those were also excluded. The final sample size determined for analysis was 4,231 requests.

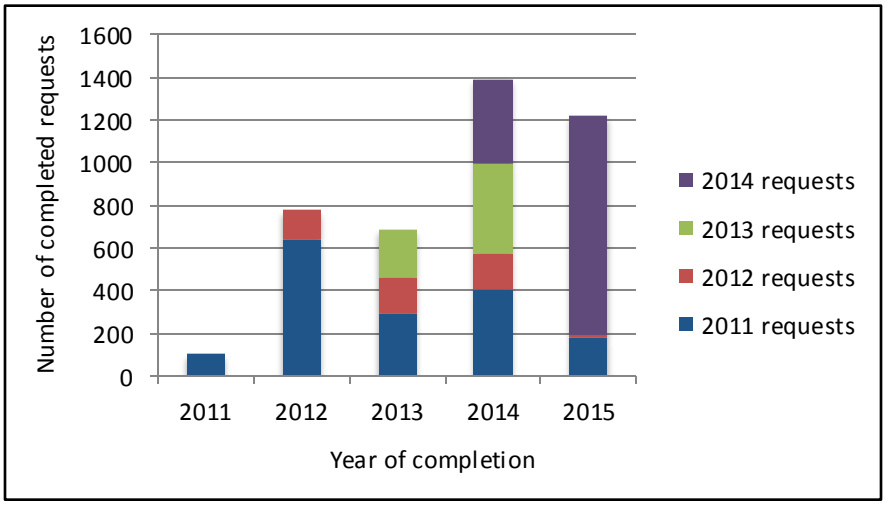


Figure 8: Completed requests by year

Based on these 4,231 requests, the number of requests completed could be determined. Figure 8 shows the number of requests completed in years 2011-2014 – the bar color corresponds to the years that the requests were originally submitted. For example, in 2012, 778 requests were completed in total – 645 of these requests were submitted in 2011 while the remaining 133 requests were submitted (and completed) in 2012. This chart gives a sense of general rate of completion over these four years. Requests submitted in 2011 seem to take longer to reach completion than those submitted in 2012, 2013, and 2014. In fact, this graph demonstrates a decrease over time in the lifecycle of completed requests. However, this chart does not show the total amount of requests submitted in those individual years for reference.

Figure 9 provides more information as to completion rates over these four years. As suggested in Figure 8, the average time for this sample of requests to go from submitted to completed decreased over time. In 2011, the average time to reach completion was 652 days, in 2012, it was 380 days, in 2013, it was 145 days, and then 128 days in 2014. Additionally, the level of variance between the requests (difference in completion times amongst the sample) decreased as well.

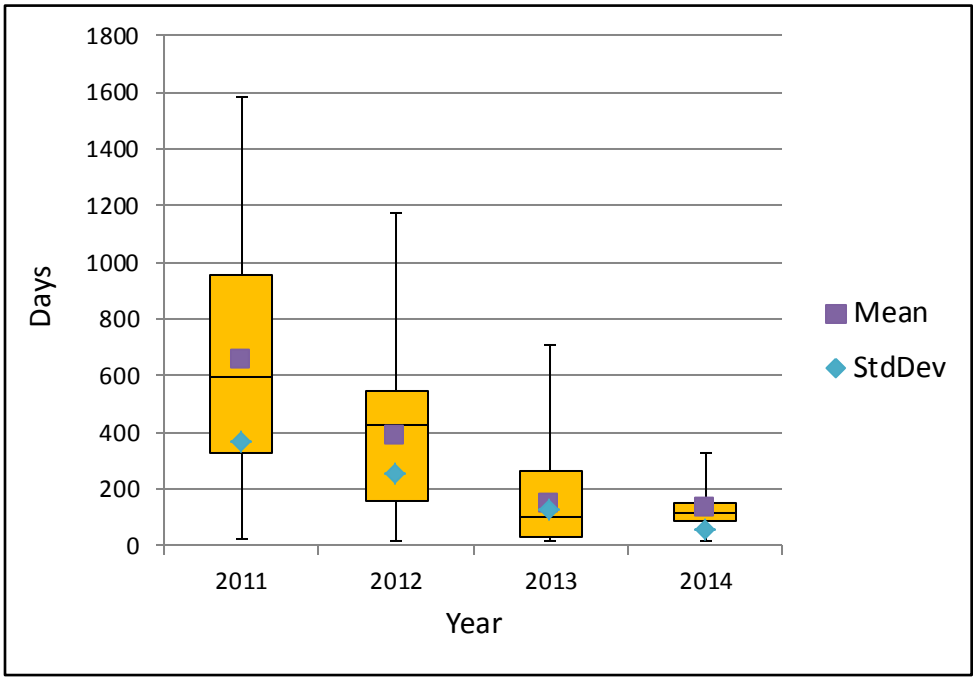


Figure 9: Completion rates over time

The decrease in average time to reach completion is an encouraging trend; potential reasons for this decrease are generally speculative. Of this sample, the largest amount of requests submitted occurred in 2011 (i.e., 1,628 requests), which also correlates with the largest amount of submissions to USCRS in general (see Figure 4). Having such a large number of requests could explain why it took so long for the 2011 sample to make their way through the UCSRS review process. However, 2014 was the second highest year of submitted requests that reached completion in this sample (i.e., 1,427), but that year has the lowest average days to reach completion. Further analyses could determine additional factors that might influence this downward trend such as number of requests completed in USCRS vs. forwarded onto IHTSDO.

Discussion

Challenges

This project encountered significant challenges throughout, primarily related to limitations of the current USCRS system and workflow. The majority of these difficulties were due to incomplete data recorded about requests in the USCRS database. For example, when attempting to determine the lifecycle of completed requests, initial analysis discovered that many completed requests in USCRS were missing the assigned SNOMED CT Concept IDs and accompanying effective times, which made identifying the *true* time of completion challenging. Additionally, further investigation of this missing data revealed that these pieces of information (i.e., SNOMED CT Concepts IDs) were manually entered rather than automatically captured when assigned. This challenge resulted in use of the UTS API, as described in the Lifecycle of completed requests section. Additionally, there are several fields in USCRS requests records that are not required. One of these fields is the "Submitted" field, which captures information about when requests were submitted to the USCRS system. While the majority of requests have this information, not all do and as a result when querying the USCRS database, it was important to remember to query certain fields (i.e., IS NULL) to capture all relevant records. A final challenge

encountered was data clean-up. The most significant difficulty came in the process of deduplicating data (i.e., USCRS requests) pulled through use of the UTS API.

Recommendations

In light of these challenges, the following recommendations are made:

- **Clean data in current USCRS system before entering into new USCRS system**

When the new USCRS system is released it is important to have this data not only so there is consistency in the information captured about requests, but also to have complete records for requests submitted to USCRS. Also, having this information would contribute toward building a corpus of data to allow for future analyses of the system. This will most likely require manual inspection of those records with missing information. As a result of the Lifecycle of completed requests analysis, 4,231 *New Concept* requests are cleaned and ready for entry into the new USCRS system; however, all other types will require data clean-up.

- **Automate the capture of assigned SNOMED CT Concept IDs**

When requests reach the "completed" end state and are entered into either the US Edition or the International version, currently the associated SNOMED CT Concept IDs are manually entered into the USCRS request records. As a result, this piece of data has not been consistently captured in request records, which makes analysis of the system difficult, but also compromises the ability for USCRS to provide a complete record of requests that are processed through the system. Consequently, this recommendation focuses on backend integration with SNOMED CT in order to automatically capture this information.

- **Conduct user assessment of USCRS to determine potential areas for providing support**

Using the information collected in this project, user analysis of USCRS can help to determine areas in which further documentation and user support can be offered. Submitter characteristics can be used to identify organizations to reach out to about user testing. Additionally, analyses, such as those done related to Rejected requests, can point to potential areas where users might benefit from guidance (e.g., tutorials). Ultimately, these types of actions work toward creating a more efficient system with a higher quality of requests.

- **Generate a calendar for edition releases and USCRS workflows**

Information pertaining to the releases of the US Edition and International version is available on USCRS, but there are additional workflows that take place around the releases that may be beneficial to publicly list. There is a quality assurance (QA) period that takes place in producing the US Edition releases – during this time, the request review process is halted in order to conduct QA and release the US Edition on time. Having this information listed publicly allows for greater transparency, but also allows for accurate capture of information in determining future analyses.

Key components of terminology request systems

Considering the challenges encountered and the resulting recommendations from this project, a list of "desired elements" or key components was developed to guide the development and enhancement of terminology request systems at NLM. Many of these elements are already possessed by USCRS – in most cases, it is just a matter of optimizing them moving forward. This list may be of use to organizations and systems outside of NLM as well. The elements listed below include features and functionality that are advantageous for supporting users, conducting analysis, and enabling future systems development.

- **Central system for external/internal requestors**

A single system for processing requests from external or internal requestors allows for consistency in processing as well as aggregation of data. In the case of USCRS, a centralized request system permits the request manager to handle all requests that come into the system for the US Extension or move them along to IHTSDO. Information about requests can be kept together and this assists in maintaining complete records for requests submitted to the system. This also allows for analyses, such those done in this project, to be conducted efficiently due to the data being located in one area.

- **Communication between requestors & request system managers**

Having a space for requestors to communicate with those evaluating requests is another important feature of a terminology request system. As demonstrated in this project, this communication allows for clarification and discussion about requests as well as captures further detail about requests. Ideally, this communication is self-contained in the system and is represented in the request record. Doing so, again, allows for consistency, but also ensures that all information about a request is contained in one location.

- **Balance between user knowledge and system QA**

Any terminology request system assumes a certain amount of knowledge from request submitters. Currently, USCRS states as a requirement that submitters should have an understanding of SNOMED CT's structure and content. Due to the specialized nature of terminologies, an assumption and requirement of user knowledge is understandable. However, there is also a role that the request system can play in supporting and optimizing the requests submitted. For example, when users submit a request for a *New Concept* in USCRS, they have to provide a parent for the concept. An example of using the system to support these requests would be to provide an auto suggest feature so when users type in a parent name, the system can pull up suggestions for potential parents that exist in SNOMED CT. USCRS has developed and begun to mobilize this feature and, looking forward, it would be useful to extend this auto suggest feature to the actual concept requests submitted.

- **Integration of terminology and request system**

Building upon the previous component, is the importance of integrating the terminology into the request system itself. In order to mobilize some of these features, such as the auto suggest, there needs to be communication between the request system and the terminology in question. Not only can this act to support the user end of the system, but backend capture of information related to requests will

also be enhanced. In the case of USCRS, full integration of the terminology and the request system can aid in capturing SNOMED CT Concept IDs as they are assigned as well as help establish the lifecycles of requests in USCRS.

- **Provide complete lifecycle view of requests**

A strong impetus in implementing many of these key components is the importance of capturing a holistic view of the lifecycle of requests. For reasons discussed in regards to previous elements, collecting comprehensive information about requests provides for strong data capture and allows for analyses to be conducted of the system and its content. Additionally, these terminology request systems are services provided to develop and enhance the terminology in question. As such, the information contained in these systems may be of interest not only to submitters and their supporting organizations or agencies, but also to organizations that own, maintain, and develop terminologies. The ability to produce complete records for requests if and when requested not only demonstrates the strength of the system itself, but enhances the relationships between the system and its users.

- **Transparent processing**

In considering terminology requests systems as services provided, it is important to provide information to users as to how requests are processed, where in the review process they are, and when they have reached an end state in the system. USCRS uses "states" to communicate within the system, and to submitters, where in the review process a request is. This information is made publicly available, see Background, but in ensuring transparency it may be of use to make a flowchart of processing available to users so they get a sense of the multiple paths a request might take. Additionally, making system generated reports available so that submitters can see the requests they have submitted and how they have made their way through the system is a useful element as well. Finally, mirroring a recommendation made earlier in this report, providing a public calendar of processing times and releases would help in tying together the external and internal processes that influence the manner in which requests are processed through a terminology request system.

Conclusion

This project provided greater insight into the requests that are included in USCRS, specifically around submitter characteristics, rejected requests, and the lifecycle of completed requests. Moving forward, the results and challenges encountered during this project can be used to further develop the USCRS system, both in the front end and backend. The project methodology can also be used as a template for future analyses of USCRS. Additionally, the suggested "key components" are intended to help build and maintain strong, effective, and efficient terminology request systems at NLM or other related organizations.

References

[1] IHTSDO. (2015). *SNOMED CT*. Retrieved from http://www.ihtsdo.org/snomed-ct.

[2] National Library of Medicine. (2011). *US Extension to SNOMED CT*. Retrieved from http://www.nlm.nih.gov/research/umls/Snomed/us_extension.html.

[3] National Library of Medicine. (2013). *US Edition of SNOMED CT*. Retrieved from http://www.nlm.nih.gov/research/umls/Snomed/us_edition.html.

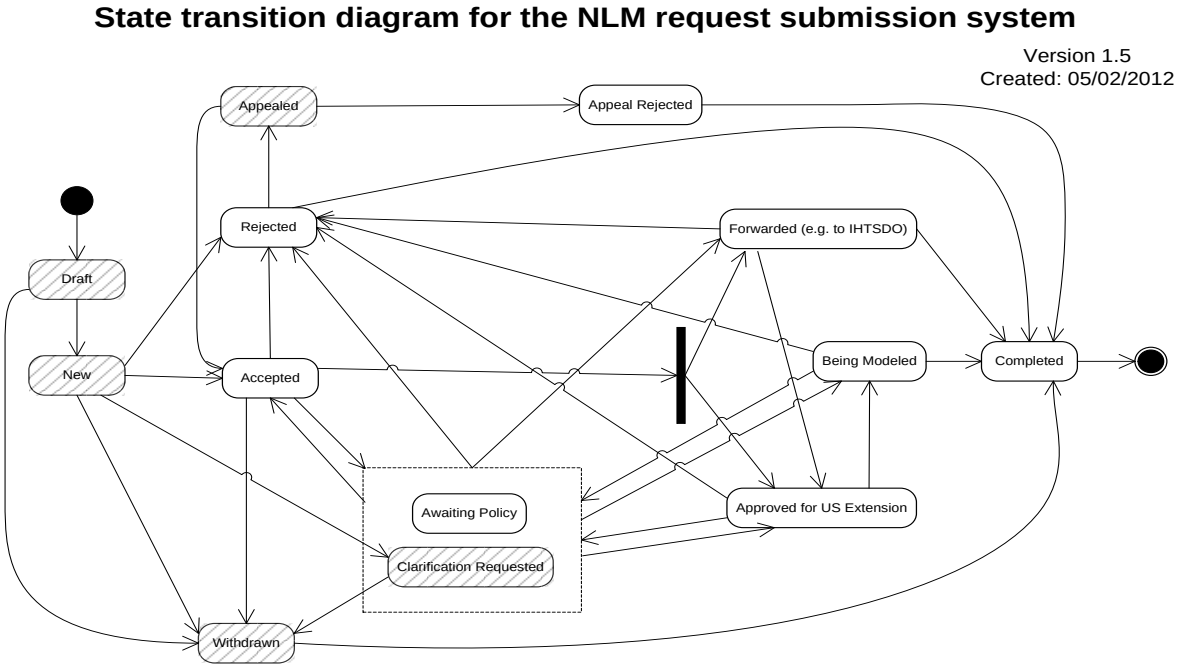
[4] National Library of Medicine. (2012). *U.S. SNOMED CT Content Request System (USCRS)*. Retrieved from https://uscrs.nlm.nih.gov/.

Acknowledgments

Thank you to Suzy Roy and Steve Emrick for acting as project sponsors and providing knowledge, support, and guidance that was instrumental to the successful completion of this project. Thank you to Dr. James Case and Sukhbir Singh for offering their expertise and guidance. Thanks to Dr. Olivier Bodenreider, Dr. James Pash, and Dr. John Kilbourne for their contributions as well. Finally, thank you to Dr. Kathel Dunn, Maureen Madden, Ariel Deardorff, Kristina Elliott, and Lori Harris for their support and advice throughout this project.

Appendices

Appendix I – USCRS state transition diagram



Appendix II – Research questions

1) Which organizations are submitting requests to USCRS?
    a. Requests by year
2) What types of requests are submitted to USCRS?
    a. Temporal breakdown of requests
3) Rejected requests
    a. What are the reasons for rejection? (i.e., evaluating Notes in USCRS for these requests)
    b. How do rejected requests breakdown by organization?
    c. How do rejected requests breakdown by type?
4) Lifespan/timeline of requests
    a. Number completed / release cycle
    b. How do requests breakdown by type?
    c. Requests approved for US Edition vs. forwarded to IHTSDO
        i. Numbers / release cycle

Appendix III – SQL queries

Submitter Characteristics:

```
SELECT submitterorganization, status, type, count(*)
FROM requests
WHERE status!='DRAFT' and submitted <='31-MAR-15'
GROUP BY submitterorganization, status, type
ORDER BY submitterorganization, count(*) desc, status
```

```
SELECT submitterorganization, status, type, count(*) FROM requests WHERE status !='DRAFT' and
submitted IS NULL GROUP BY submitterorganization, status, type ORDER BY submitterorganization,
count(*) desc, status;
```

Rejected requests & notes:

```
select *
from requests
full outer join request_notes
on requests.id=request_notes.request_id
where requests.status='REJECTED';
```

Lifecycle requests:

```
select requests.id, concepts.terminologyid, requests.submitterorganization,
newconcept_requests.concept_id, descriptions.term, descriptions.type, requests.submitted from
concepts
join newconcept_requests on concepts.id = newconcept_requests.concept_id
join descriptions on newconcept_requests.concept_id=descriptions.concept_id
join requests on newconcept_requests.id=requests.id
where requests.status='COMPLETED' and requests.submitted < '19-FEB-15' and descriptions.type='1'
and requests.type = 'newconcept';
```

** The date of February 19, 2015 was used for this query because that was when QA began for the March 2015 release, which meant a pause in request processing. This date was used in order to determine the most accurate sample in light of the current workflow.

Appendix IV – Full list of submitter organizations

Submitting organizations	Type of organization	Number of requests
3M Health Information Systems, Inc.	Company	164
Academy of Nutrition and Dietetics	Professional organization	348
American Dental Association (ADA)	Professional organization	1
American Medical Association (AMA)	Professional organization	314
American Society of Clinical Oncology (ASCO)	Professional organization	13
Apelon, Inc.	Company	95
Association of periOperative Registered Nurses (AORN)	Professional organization	60
Association of Public Health Laboratories (APHL)	Professional organization	479
Bassett Healthcare	Company	1
Canada Health Infoway	Company	3
CareCloud	Company	4
Case Western Reserve University	Academic institution	58
Centers for Disease Control	Government agency	453
Cerner Corporation	Company	562
Chidren's Hospital Boston	Company	494
Children's Hospital and Clinics of Minneapolis	Company	1
Cleveland Clinic	Company	1
Clinical Informatics Consulting	Company (consulting)	426
College of American Pathologists	Company	6
Columbia University	Academic institution	136
Data Networks Corporation	Company (consulting)	65
Department of Veteran's Affairs (VA)	Government agency	172
Digital Imaging and Communications in Medicine (DICOM) Standards Committee	Committee	11
Epic System Corporation	Company	1
First DataBank	Company	12
Food and Drug Administration (FDA)	Government agency	66
Frank McKinney Group LLC	Company (consulting)	1
General Dynamics IT	Company	14
Harvard University	Academic institution	320
Health Catalyst	Company	3
Health Language	Company	1
Hospital Sisters Health System	Company	1

IBM	Company	3
IHTSDO	Company	3
Independent Consultant	Company (consulting)	80
Intelligent Medical Objects (IMO)	Company	4
Intermountain Healthcare	Company	545
Intermountain Healthcare/GE Healthcare	Collaborative	81
Kaiser Permanente	Company	5,961
Lantana Consulting Group	Company (consulting)	8
LMI	Company (consulting)	4
Mayo Clinic	Company	15
Mays Systems	Company	468
McKesson Specialty Care Solutions	Company	61
MD Partners, Inc.	Company (consulting)	21
Medhost	Company	1
MITRE Corporation	Company	7
National Committee for Quality Assurance	Company	78
National Library of Medicine	Government agency	254
New Jersey Institute of Technology (NJIT)	Academic institution	178
New York State	State agency	205
NIH Clinical Center	Government agency	85
Oklahoma Foundation for Medical Quality (OFMQ)	Company	5
Partners Healthcare	Company	30
Pharmacy e-HIT Collaborative	Collaborative	286
Quality Insights of Pennsylvania	Company	12
Stanford University	Academic institution	198
Statistical Keys for Medical Landscape (skml.fr)	Company	1
Swedish Covenant Hospital	Company	2
Tennessee Department of Health	State agency	1
University of California, Davis	Academic institution	5
University of Colorado Denver	Academic institution	51
University of Nebraska	Academic institution	2,739
University of Utah	Academic institution	38
Virginia Tech	Academic institution	24
Warren Associates, LLC	Company (consulting)	1
West Coast Informatics, LLC	Company (consulting)	4
Unknown	Unknown	361