Conducting Taxonomy Validation: Healthcare Example

Joseph A. Busch, Taxonomy Strategies, Washington, D.C., USA

Background

Taxonomy Strategies spent almost a year developing a knowledge organization system (KOS) for the Centers for Medicare and Medicaid Services (CMS) to support consumers in making better health care decisions. Our starting point for the Consumer Health Care Taxonomy, was consumers (including but not limited to beneficiaries) or caregiver looking for health care information and services. The Consumer Health Care Taxonomy was designed from the outset to support the types of queries a consumer health care information service such as a website might get from a wide variety of consumers in a wide variety of care conditions. While the consumer is the primary audience, a consumer health care website and its taxonomy exist in an ecosystem of other stakeholders and individuals expert in the needs of consumers and CMS's systems.

This paper provides background on the existing CMS Medicare.gov datasets, an overview of the Consumer Health Care Taxonomy, and then discusses the methods used to refine the taxonomy.

Background

CMS currently maintains six websites which enable consumers to search for health care service providers or suppliers based on their postal code, and in some cases with additional criteria such as physicians by gender or body part treated. These websites cover hospitals[1], nursing homes[2], physicians[3], home health care services[4], dialysis facilities[5] and medical suppliers[6]. CMS provides separate search interfaces to separate Medicare.gov datasets[7]. Most Compare websites provide access only by zip code and name of the service provider. Shown in Figure 1, the Physician Compare website also offers additional criteria to help identify physicians by name, specialty areas, and other ways.

[1] http://www.medicare.gov/hospitalcompare/.
[2] http://www.medicare.gov/NursingHomeCompare/.
[3] http://www.medicare.gov/physiciancompare/.
[4] http://www.medicare.gov/homehealthcompare/.
[5] http://www.medicare.gov/Dialysisfacilitycompare/.
[6] http://www.medicare.gov/supplierdirectory/.
[7] https://data.medicare.gov/.

Figure 1-Physician compare website showing some of the search methods.

Each of the datasets is produced independently with little or no standardization in the data structure and data values. For example, unique identifiers for hospitals might be assigned based on an individual facility, or for a whole system; or categories of services might be identified by a column heading rather than an explicit human-searchable data value. Figure 2 summarizes our inventory of the Medicare.gov datasets. There are differences across the Medicare.gov datasets. Physician Compare and Supplier Directory datasets, for example, were consolidated into a single table which included all the information available for each provider, while the other datasets had multiple tables which contained provider information. Hospital Compare had the largest number of files (59) in their dataset. More Compare websites are being developed and launched that cover additional care settings.

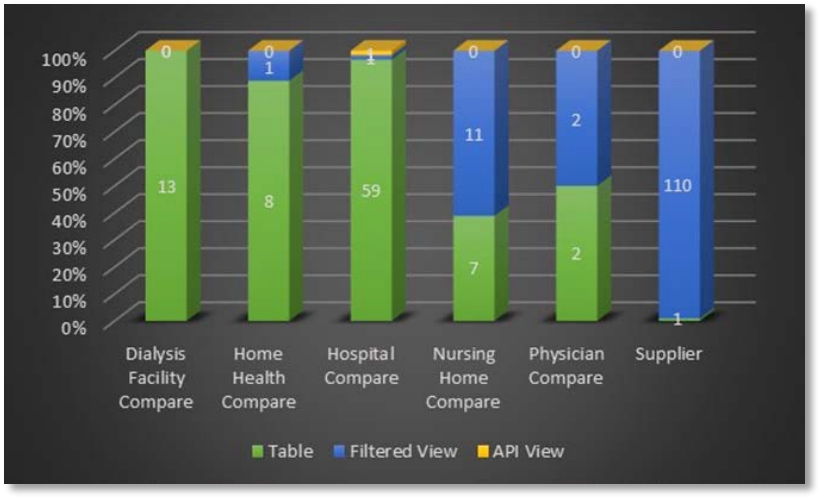

Figure 2-Number of tables in each Medicare.gov dataset on data.medicare.gov.

The Consumer Health Care Taxonomy

The Consumer Health Care Taxonomy needs to function as middleware that translates consumer queries into the language necessary for retrieval of data from Medicare.gov datasets and Good to Know (GTK) content. The Consumer Health Care Taxonomy should:

- Provide enough information for any user, tool, or program to find and use content in any Medicare.gov dataset or GTK content.
- Define what vocabularies are needed to support consumer health care decision making.
- Identify authoritative vocabulary sources for each taxonomy facet.
- Provide vocabularies for each taxonomy facet that are sufficiently defined to be used to build a functional application.
- Be readily extensible to support new application requirements.
- Be flexible enough to accommodate additions of missing categories and changes to existing categories as needed.
- Define relationships between the vocabularies useful for searching Medicare.gov datasets and GTK content.

The Consumer Health Care Taxonomy is currently a collection of eleven facets or vocabularies. Each facet is comprised of entry terms, synonyms, quasi-synonyms, hierarchical relationships inside the facet and relationships across the facets. The purpose of this project was to develop this framework with enough terms and relationships to be effective, but it is not exhaustive. As the Taxonomy evolves, more entry terms, synonyms, and relationships will be added as the ultimate application design is determined and the behavior of consumers on the site is revealed. The facets were identified through interviews, research and analysis as discrete conceptual areas important to consumer health care decision-making search paths and to surfacing GTK content. Figure 3 shows the eleven facets in the Consumer Health Care Taxonomy.

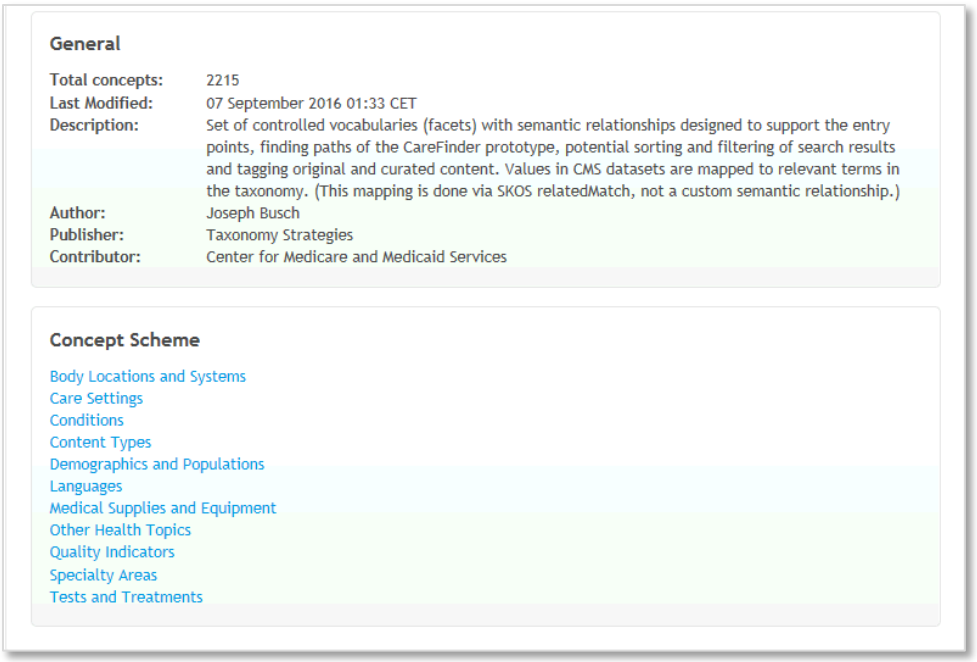

Figure 3-Eleven facets in the Consumer Health Care Taxonomy displayed in the PoolParty Linked Data frontend.

The real power of the Consumer Health Care Taxonomy is the relationships between terms in different facets. These relationships provide the mechanism for a consumer searching by the name of a condition to find a physician or a care setting specializing in that condition. The relationships help the search engine identify relevant Medicare.gov dataset information and GTK content related to the consumer's query. Only relationships that are needed to support consumer health care decision-making have been created. Currently, six of the eleven facets have one or more relationships to other facets. Table 1 lists the current relationships between facets. These relationships are in pairs and each has a semantic label providing more information on how the facets (and terms in the facets) are related.

Facet Class	Relation → ← Inverse relation	Facet Class
Body Locations and Systems	is affected by → ← affects body location	Conditions
Conditions	has treatment of → ← is treatment for	Tests & Treatments
Conditions	is concern of → ← Is concerned about	Specialty areas
Conditions	needs medical supply → ← is needed for condition	Medical Equipment & Supplies
Care Setting	is location for treatment → ← is treatment provided in	Tests & Treatments
Care Settings	specializes in → ← is specialty of	Specialty Areas
Medical Supplies & Equipment	is used in treatment → ← uses medical supply	Tests & Treatments
Specialty Areas	includes treatment of →	Tests & Treatments

Facet Class	Relation → ← Inverse relation	Facet Class
	← is part of practice area	
Care Settings	has focus of condition à ← is focused on in setting	Conditions
Body Locations & Systems	location is treated by à ← treats body location	Tests & Treatments

Table 1-Current relationships between facets in the Consumer Health Care Taxonomy.

Table 2 provides an example of the semantic relationships for the condition "End-stage renal disease".

Concept	Relation → ← Inverse relation	Concept
Kidneys	is affected by → ← affects body location	End-stage renal disease
End-stage renal disease	has treatment of → ← is treatment for	Dialysis
End-stage renal disease	is concern of → ← Is concerned about	Nephrology
End-stage renal disease	needs medical supply → ← is needed for condition	Dialysis Equipment & Supplies
Dialysis Facilities	is location for treatment → ← is treatment provided in	Dialysis
Dialysis Facilities	specializes in → ← is specialty of	Dialysis Services
Dialysis Equipment & Supplies	is used in treatment → ← uses medical supply	Dialysis
Nephrology	includes treatment of → ← is part of practice area	Dialysis
Dialysis Facilities	has focus of condition → ← is focused on in setting	End-stage renal disease
Kidneys	location is treated by → ← treats body location	Dialysis

Table 2-Semantic relationships for the condition "End-stage renal disease".

Taxonomy Validation Methods

The taxonomy team constructed several mechanisms that helped validate both data retrieval from Medicare.gov datasets and so-called "Good to Know" web content from other sources. The methods used to refine the taxonomy included:

- Using expert reviewers,
- Gathering query logs,
- Collecting user stories,
- Walk-through taxonomy validation,
- Building simple validation tool,
- Validating each version of the taxonomy,
- Using relevant content, and

- Noting changes needed.

Expert reviewers

Best practice. Consult subject matter experts (SME's) throughout the Taxonomy development process. Get their comments and questions about the structure and content of the Taxonomy as it evolves.

Throughout the Taxonomy development process, we consulted with subject matter experts (SME's) to get their comments and questions about the structure and content of the Consumer Health Care Taxonomy as it evolved.

- For the initial set of facets, we asked the CMS Web and New Media Group (WNMG) contacts to review the set of discrete conceptual areas as a reasonable starting point to build out the taxonomy in three iterations.
- For the V.1 Taxonomy, we asked the NORC at the University of Chicago, an independent research organization, and the Center for Clinical Standards and Quality (CCSQ) team, the Compare website owners, to review the taxonomy facet labels; and we requested the WNMG contacts to review all the concepts in the top levels of the taxonomy.
- For the V.2 Taxonomy, we asked NORC to identify several SME's to review the relationships between the key taxonomy facets Conditions, Treatments and Specialties, and we requested the core CCSQ team to review the overall structure and browse all the concepts.

We also worked with NORC to work with the CCSQ team to review the Quality Indicators framework we developed, and to map our inventory of Quality Measures to the proposed Quality Indicators framework.

Query logs

Best practice. Gather query logs to identify most common terms used to search for related content. Ongoing, monitor query logs to identify popular and emerging concepts and new relationships.

We analyzed the Physician Compare and Medline Plus query logs to identify the most common terms that users searched on. These terms were consolidated around similar concepts and grouped by type such as condition, treatment, drug, etc. This early validation method were helpful in evaluating the basic taxonomy framework of discrete facets with semantic relationships between them, plus concepts with synonyms including non-technical labels, abbreviations and acronyms.

User stories

Best practice. Collect user stories from stakeholders. Use the stories to walk through how the taxonomy enables content findability.

Once the validation approach was agreed, we gathered user stories from the tool owners at CMS, NORC, and the WNMG contacts. These user stories were written as anonymized narratives of real people looking for help with their health needs.

Each narrative was distilled into "searchable" components. "Claire's" story, for instance, explores the many problems she has had with her new hip, but the story distills into a search for orthopedic surgeons and a rehab facility. Simpler user stories typically narrow to a single search, more complicated stories – like Claire's – lead to multiple searches.

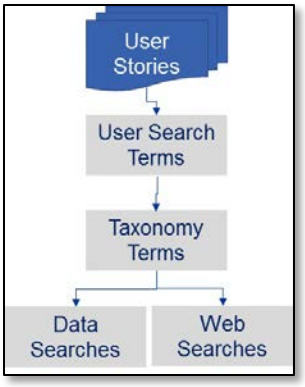

Figure 4-Each narrative was distilled into searchable components, and then translated through the Taxonomy into data and web searches.

Taxonomy walk-throughs

Best practice. Walk-through taxonomy interaction process.

Figure 4 shows the flow of the search terms from the user stories that are used as input into the Consumer Health Care Taxonomy to retrieve the entry term, synonym, semantically related terms, and CMS dataset values. Figure 5 provides an example of how a specific user story is transformed through the Taxonomy into a Compare dataset search and web search.

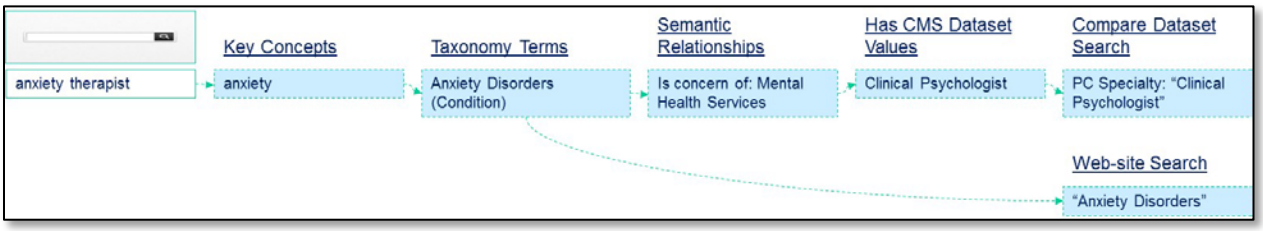

Figure 5-Taxonomy walk-through of user story: "I have a history of anxiety. Find a therapist."

Simple validation tool

Best practice. If feasible, develop a simple application to automate the validation process.

For validating the V.1 and V.2 Taxonomy we used data.medicare.gov to retrieve data from the Compare Sites. For the validation of the V.3 Taxonomy we used a custom-built simple validation tool that simplifies searches of the same data sources as illustrated in Figure 6. The benefits of the tool include being able to search multiple datasets at one time, data field searches instead of simple text string searches, sorting search results by star ratings (when available), keeping the taxonomy in sync with the dataset values, and the ability to save search results to a file as illustrated in Figure 7.

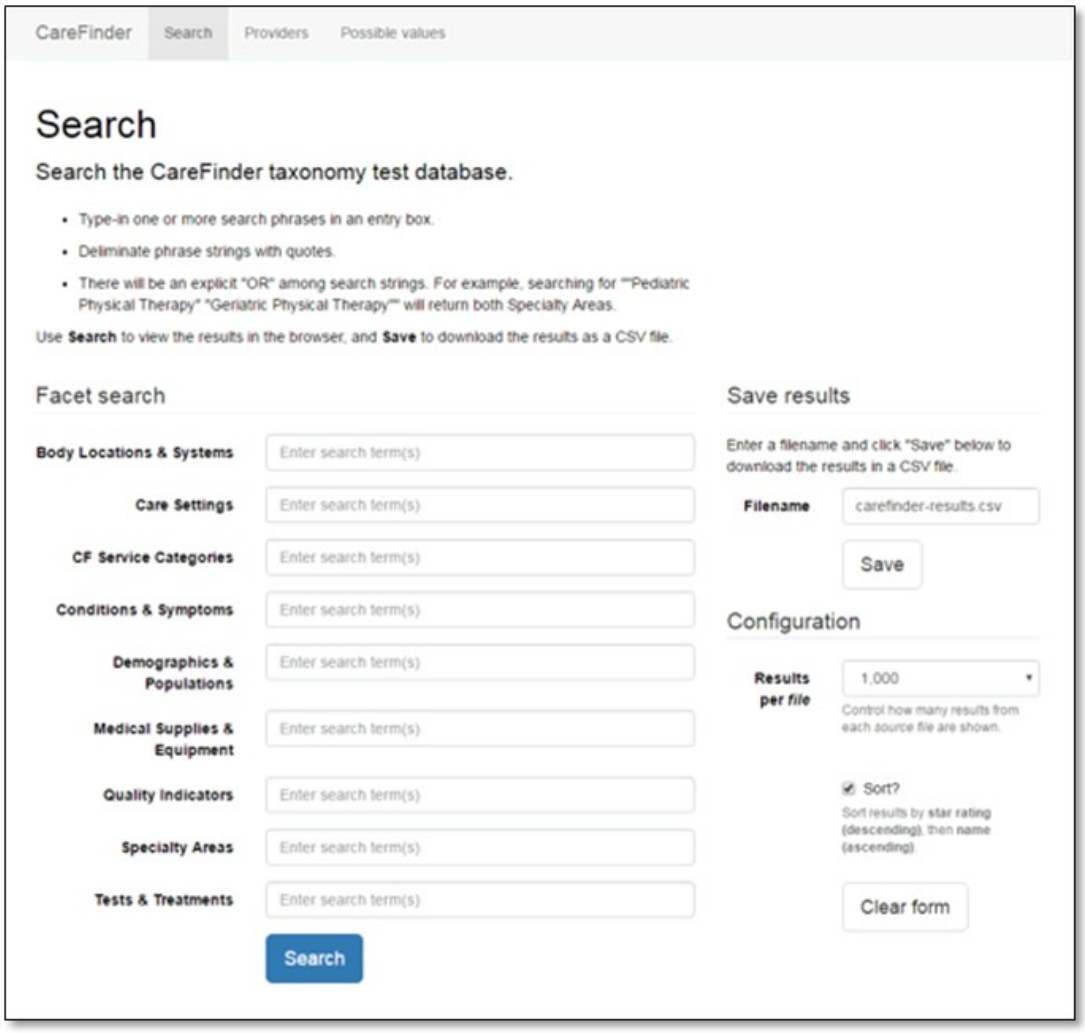

Figure 6-Screen shot of custom-developed validation tool for searching across CMS Compare datasets.

A	B	C	D	E	AD	AQ	AR	AS	AT	AU	CE	CL	CM	CN
Query	Specialty Areas: clinical psychologist													
Files	Physician Compare													
Count	10	58112												
Provider ID	Name	Zip	Star rating	Source	Gender	Claims based hospital affiliation LBN 1	Claims ba	Claims ba	Claims ba	Claims ba	Primary specialty	Secondary	Secondary	Secon
"1225000979"	AALBERS, CAROL J	897034269		Physician Compare	F						CLINICAL PSYCHOLOGIST			
"1225000979"	AALBERS, CAROL J	897034269		Physician Compare	F						CLINICAL PSYCHOLOGIST			
"1639108095"	AARON, LESLIE A	982252914		Physician Compare	F						CLINICAL PSYCHOLOGIST			
"1639108095"	AARON, LESLIE A	982252914		Physician Compare	F						CLINICAL PSYCHOLOGIST			
"1821010695"	AARON, LISA A	150909213		Physician Compare	F						CLINICAL PSYCHOLOGIST			
"1821010695"	AARON, LISA A	150909213		Physician Compare	F						CLINICAL PSYCHOLOGIST			
"1851501787"	AARONSON, BARRY L	926124694		Physician Compare	M	HOAG MEMORIAL HOSPITAL PRESBYTERIAN					CLINICAL PSYCHOLOGIST			
"1851501787"	AARONSON, BARRY L	926124694		Physician Compare	M	HOAG MEMORIAL HOSPITAL PRESBYTERIAN					CLINICAL PSYCHOLOGIST			
"1528113958"	AARSVOLD, RAMONA L	787598444		Physician Compare	F						CLINICAL PSYCHOLOGIST			

Figure 7-File of saved CMS Compare dataset search results generated by the validation tool.

Iterative validation

Best practice. Do iterative validation throughout the taxonomy development process.

We performed our validation tests on each of the three revisions of the Consumer Health Care Taxonomy. In each iteration, we used a random selection of use cases, then used the taxonomy to

identify the best search terms for CMS datasets – these would be CMS dataset values, and the best search terms for Good to Know content – these would be taxonomy terms.

- For the validation of V.1 Taxonomy the sematic relationships between facets were not completed so we were only able to retrieve the entry term and synonym for a particular user search term.
- For V.2 validation some relationships between facets were developed and some of the associated terms were retrieved.
- V.3 validation reaffirmed the structure of the V.3 Taxonomy, which has all semantic relationships built out (but not all were necessarily populated), and the taxonomy's ability to find Good to Know content. Both V.2 and V.3 validation exercises were "blind tests," meaning the user stories were not shared with the Taxonomy development team.

In each revision we were able to retrieve all related terms identified for a particular user search term. In other words, the taxonomy delivered more – and more relevant – information than specifically asked for by the user, thereby helping him or her make better healthcare decisions. The semantically related terms helps us retrieve the additionally relevant information.

Use of relevant content

Best practice. Use relevant content to make the validation process compelling.

While the Medicare.gov datasets are fairly well understood, Good to Know content sources are not entirely clear. After consulting with CCSQ and WNMG we knew we should be searching only authoritative governmental sources. We evaluated several candidates before settling on Medline Plus as the target for validating the good to know content.

As with the validation against Medicare.gov datasets, we used the search terms from the user stories as inputs into the Good to Know queries to Medline Plus and compared the results to searches using Taxonomy terms. Figure 8 shows a side-by-side comparison of user terms vs. Taxonomy terms for the user story: "I have a history of anxiety. Find a therapist." Figure 9 shows the Good to Know content relevance scoring worksheet. The top six Medline Plus hits were shown to three members of the Taxonomy team to make a personal determination of how helpful the article would be if he or she were the user looking for information. If all articles were relevant the search would receive a score of 6; if only five articles were relevant the search would receive a score of 5, and so forth. The average relevance scores (averaging the scores of the tribunal) were compared to the relevance scores of the Taxonomy term to determine which entry term would yield more useful results for a user. Overall, with few exceptions, the Taxonomy term yielded more results that were useful as determined by our team.

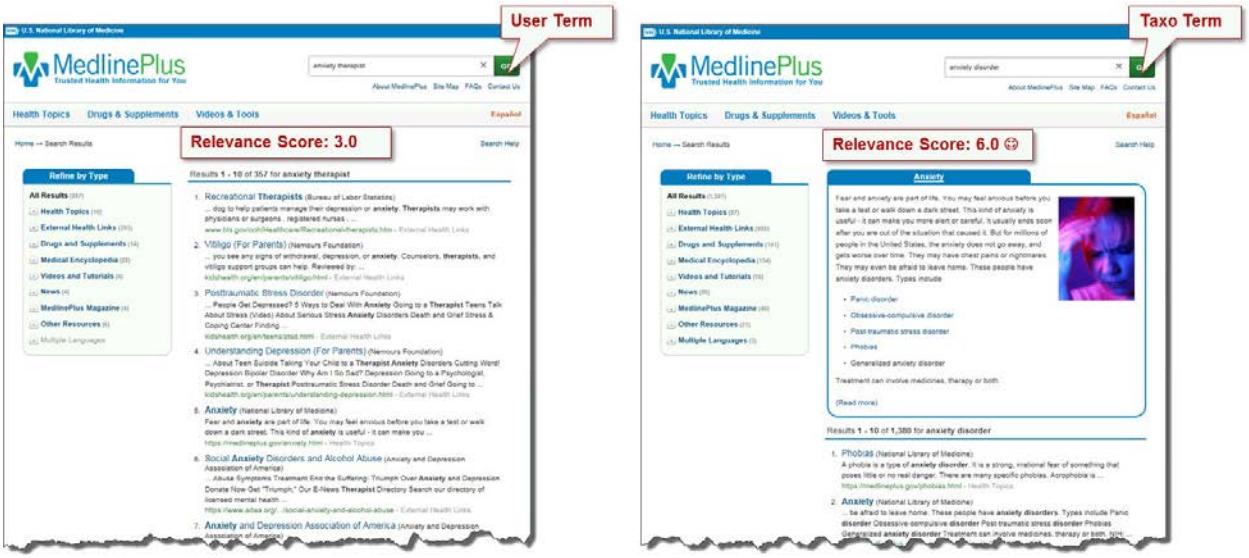

Figure 8-Side-by-side comparison of user term vs. Taxonomy term for the user story: "I have a history of anxiety. Find a therapist."

	Source ID	Simplified user story	User Search	JB	AS	ML	AVG	Taxonomy Search	JB	AS	ML	AVG	Score
2	8	Find surgeons who perform hip replacement surgeries	Hip replacement surgery	5	5	5	5.00	Hip replacement	6	6	6	6.00	
3	9	Identify rehab facilities	Rehab facilities	3	2	4	3.00	Rehabilitation facilities	4	3	4	3.67	
4	10	Identify who can provide physical therapy services in the home	Home physical therapy	2	2	2	2.00	Home health care	5	6	5	5.33	
5	13	I have a history of anxiety. Find a therapist.	Anxiety therapist	2	2	5	3.00	Anxiety Disorder	6	6	6	6.00	
6	31	Find dialysis close to home	Dialysis facilities	6	5	6	5.67	Dialysis services	6	6	6	6.00	
7	32	Where to get physical therapy for a broken leg	Broken leg	2	3	3	2.67	Leg fractures	3	3	2	2.67	
8			Physical therapy	2	2	1	1.67	Physical therapies	3	2	2	2.33	
9	49	Find a specialist to assess a seniors memory loss	Memory loss specialist	2	3	3	2.67	Dementia	3	2	4	3.00	
10	66	How do low levels of thyroid stimulating hormone impact early fetal development?	Thyroid pregnancy	3	3	3	3.00	Thyroid Diseases Pregnancy	5	3	5	4.33	
11	72	Find doctor that handles elbow fractures	Doctor for broken elbow	2	2	3	2.33	Elbow Fractures	2	3	3	2.67	
12													
13	How to score:												
14	Identify the number of relevant, i.e., good to know items from the top 6, and enter that number for each Medline Plus user search and each Medline Plus taxonomy search												
15													
16		Equal or Better than user search											
17		Worse than user search											

Figure 9-Good to Know content relevance scoring worksheet.

Note changes needed

Best practice. Log changes identified in the validation process. Then, prioritize changes based on their impact and process them as part of the next taxonomy iteration.

Issues found in the V.1 and V.2 validations were addressed in the subsequent taxonomy version. Issues found in the V.3 validation as well as any outstanding issues were entered into a Taxonomy Change Log. Disposition of these issues will happen in subsequent project phases. Further validation and evolution of the taxonomy will continue with the use of web analytics such as query logs once the site is live. This will help identify what users are searching for, and if more terms, or relationships should be added to the taxonomy.

Conclusions

Taxonomy validation helps the client visualize the taxonomy. This taxonomy is middleware, it's not obviously seen in the user interface. The taxonomy is invisible in the application prototype, it's hard to visualize how it works.

Project sponsors need to ask questions like – How do I know that the Taxonomy is well-designed? Is the Taxonomy on the right track? Is it going to perform the functions that we need in our application? The goal of Taxonomy validation is to respond to these questions.

In the early stages of development, Taxonomy validation helps to clarify and refine the structure and concepts in the baseline taxonomy. In later stages of development, Taxonomy validation simulates how the taxonomy will be used and shows how it will perform in those use cases. This can help to illustrate and refine the functional requirements for application developers. Throughout the Taxonomy development process, the overall goal of validation is to create confidence that the Taxonomy will perform as required when it is deployed. Validation is a confidence building activity that requires engagement with project sponsors throughout the process.