

KOS Design for Health Care Decision-making Based on Consumer Criteria and User Stories

Joseph A. Busch, Taxonomy Strategies, Washington, D.C., USA
Vivian Bliss, Kirkland, WA, USA

Background

Taxonomy Strategies spent almost a year developing a knowledge organization system (KOS) to support consumers in making better health care decisions. Our starting point for the Consumer Health Care Taxonomy, was the beneficiary 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' systems.

From an information science perspective, this is an information seeking problem. Elisabeth Davenport's 2010 ARIST chapter¹ provides a recent critique of this long standing (of more than 30 years) research theme. In particular she critiques methods that characterize qualitative social science research including critical incident technique, focus groups and micro-moment time-line interview in sense-making methodology. To Davenport's critique I would add that many information seeking studies have involved students in an academic environment, or engineers and scientists in research laboratory environments. Davenport calls this "scientificity". Consumer decision-making seems intrinsically different from scientific research, and certain consumer decisions such as health care decisions are particularly made in the context of life events. Most consumers search for health information on the Internet, usually starting with an organic search engine. The most commonly researched topics are diseases or conditions, treatments or procedures, and doctors or other health professionals. Half of online health information research is on behalf of someone else.²

From a health care perspective, this is a quality of care problem. The discourse around quality of care has been defined around two poles—evidence-based decision-making by clinicians vs. factors that patients identify as most important such as cost, qualifications and accessibility of care.³ While data about the quality of care is required to be gathered by hospitals and many other health care settings, there is evidence that patient narratives are of more interest to consumers, and easier for them to

¹ E. Davenport. "Confessional Methods and Everyday Life Information Seeking." *44 Annual Review of Information Science & Technology* (2010) pp.533-562.

² Pew Research Center. "Health Fact Sheet." (December 16, 2013). <http://www.pewinternet.org/fact-sheets/health-fact-sheet/>. Last checked October 3, 2016.

³ J. Hibbard, S. Sofaer. "Best Practices in Public Reporting No. 1: How to Effectively Present Health Care Performance Data to Consumers." AHRQ Publication No. 10-0082-EF. June 2010. <http://archive.ahrq.gov/professionals/quality-patient-safety/quality-resources/tools/pubrptguide1/pubrptguide1.pdf>. Last checked October 3, 2016.

understand.⁴ It's notable that Yelp now includes reviews of hospitals, doctors and other health care settings.

From a KOS development perspective, most health care KOS were originally designed to support researchers, clinicians and health insurers. We believe that these KOS are useful sources to build consumer-oriented health care KOS, rather than starting from scratch. Similarly, consumer terminology used in health care related activities can be useful to improving existing health care KOS. For example, Researchers at the University of Utah Department of Bioinformatics identified unique terms in PatientsLikeMe.com⁵ webpages and mapped them to the Unified Medical Language System (UMLS) Metathesaurus⁶ to identify candidate terms.⁷ It's interesting to note that most adults in the U.S. keep track of at least one health indicator.⁸

Consumer Health Care Taxonomy Sources

This section reviews the sources from which we drew to create the Consumer Health Care Taxonomy.

- CareFinder Prototype.
- Medicare "Compare" websites and datasets.
- Existing U.S. Centers for Medicare and Medicaid Services (CMS) and the U.S. Department of Health and Human Services (HHS) taxonomies.
- Authoritative sources, websites and query logs.
- Interviews with SMEs and stakeholders.
- User stories.

CMS Prototype

The U.S. Centers for Medicare and Medicaid Services (CMS) have developed a health care website prototype⁹ that envisions a new way for the to leverage their massive datasets to support consumers in making better health care decisions. The website would engage beneficiaries, their caregivers, medical professionals, and the general public to help all consumers get the best health care at the best price. The website prototype shows how the site should work, and a design guide describes both the rationale behind design decisions and some of the "how to" of designing certain features.

Medicare "Compare" Websites and Datasets

CMS currently maintains six Medicare websites which enables consumers to search for health care service providers or suppliers based on their postal code and some cases with additional criteria such as

⁴ M. Schlesinger, R. Grob, D. Shaller, S. C. Martino, A. M. Parker, M. L. Finucane, J. L. Cerully, L. Rybowski "Taking Patients' Narratives about Clinicians from Anecdote to Science." 373 *New England Journal of Medicine* (August 13, 2015) pp. 675-679. <http://www.nejm.org/doi/full/10.1056/NEJMs1502361>. Last checked: September 7, 2016.

⁵ <http://www.PatientsLikeMe.com/> is a free website where people can share their health data to track their progress.

⁶ <https://www.nlm.nih.gov/research/umls/>.

⁷ K. M. Doing-Harris, Q. Zeng-Treitler. "Computer-Assisted Update of a Consumer Health Vocabulary through Mining of Social Network Data." 13(2) *J Med Internet Res* (2011) p. e37.

<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3221384/>. Last checked: September 7, 2016.

⁸ Pew Research Center. "Tracking for Health." (January 28, 2013).

<http://www.pewinternet.org/2013/01/28/tracking-for-health/>. Last checked October 3, 2016.

⁹ <http://qa.assets.cms.gov/resources/framework/carefinder/>.

physicians by gender or body part treated. These websites cover hospitals¹⁰, nursing homes¹¹, physicians¹², home health care services¹³, dialysis facilities¹⁴ and medical suppliers¹⁵.

Taxonomy vs. data mapping project. At the beginning of this work we resolved whether we were going to build a consumer health care taxonomy to solve the problem of mapping variations in Medicare.gov Compare dataset field values; or we were going to build a consumer health care taxonomy to solve the problem of translating consumer searches into searches that would retrieve relevant entries from Medicare.gov datasets, as well as relevant “Good to Know” (GTK) content from other sources such as health care-focused websites. Together with CMS, we decided that the consumer health care taxonomy needed to be primarily concerned with interpreting consumer searches.

Inventories. We collected and analyzed the data dictionaries for each of the Medicare.gov datasets¹⁶ to identify and evaluate the controlled vocabularies currently in use. In the first pass, we were concerned that each of the Medicare.gov datasets was 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.

We also conducted an inventory of the downloadable Medicare.gov datasets available on data.medicare.gov summarized in Figure 1. In this process we discovered that again there were 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.

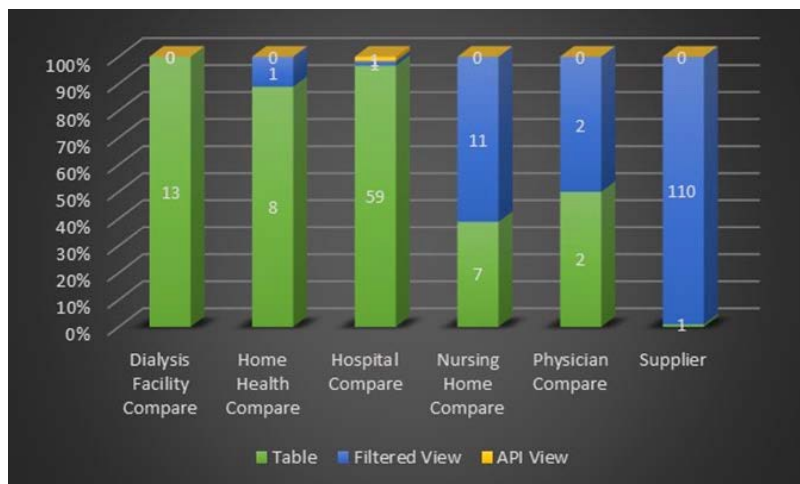


Figure 1-Number of tables in each Medicare.gov dataset on data.medicare.gov.

¹⁰ <http://www.medicare.gov/hospitalcompare/>.

¹¹ <http://www.medicare.gov/NursingHomeCompare/>.

¹² <http://www.medicare.gov/physiciancompare/>.

¹³ <http://www.medicare.gov/homehealthcompare/>.

¹⁴ <http://www.medicare.gov/Dialysisfacilitycompare/>.

¹⁵ <http://www.medicare.gov/supplierdirectory/>.

¹⁶ <https://data.medicare.gov/>.

We also developed a comprehensive list of all Quality Measures from Compare websites as well as Center for Clinical Standards and Quality (CCSQ) areas which do not have Compare websites. The goal of this exercise was not only to capture all the measures, but also to compare the types of data used in each Compare site, and consider how measures do or don't support the search needs of consumers and suggest what could be done about that. This inventory was later used to inform the development of the Quality Indicators taxonomy framework.

Existing CMS and HHS taxonomies

Taxonomy Strategies has previous experience working with U.S. Department of Health and Human Services (HHS) agencies including Agency for Health Research and Quality (AHRQ), CMS, Food and Drug Administration (FDA) and Substance Abuse and Mental Health Services Administration (SAMHSA) to create working taxonomies. In particular we leveraged past work with the SAMHSA and Healthcare.gov which is administered by CMS. The SAMHSA store¹⁷ taxonomy was intended to help health care professionals and families find and order (or download) publications relevant to their needs. The Healthcare.gov taxonomy¹⁸ was intended to organize and present general consumer-facing content about health insurance, to answer consumer questions, and to help customer support services as they were developed. In these earlier health care taxonomy development projects we developed frameworks that identified key facets, such as Conditions and Treatments that identified the key contexts for finding relevant content. We also validated their effectiveness with usability activities such as sorting, retrieving, and tagging content to measure how complete and consistent the results were. The methods and learnings from these earlier projects helped to inform the Consumer Health Care Taxonomy approach.

Authoritative sources, websites and query logs

We referred to and studied numerous websites and many authoritative resources. CMS and other U.S. federal agencies have conducted research on how to improve health care quality, as well as information to support consumer decision-making. These are documented in the bibliography of Consumer Health Care Background Resources, which includes data dictionaries, vocabularies, internal reports, government reports, government websites, Compare web sites, Foresee surveys, external reports, articles, websites, and more.

We identified and documented more than 100 vocabulary sources for developing the initial Consumer Health Care Taxonomy concepts including:

- The National Library of Medicine's Medical Subject Headings (known as MeSH), a thesaurus for indexing bio-medical books and articles;¹⁹
- The International Classification of Diseases (ICD), which is used by CMS and other health insurance providers to classify diseases and conditions;²⁰ and
- The Unified Medical Language System (UMLS), a mapping of more than 100 vocabularies and classification systems including MeSH and ICD.²¹

¹⁷ <http://store.samhsa.gov/>.

¹⁸ <http://www.taxonomystrategies.com/html/CMSFFE/>.

¹⁹ <https://www.nlm.nih.gov/mesh/>.

²⁰ <http://www.who.int/classifications/icd/>.

²¹ <https://www.nlm.nih.gov/research/umls/>.

We analyzed query logs from Physician Compare and MedLine Plus²² to help identify unique facets, relevant terms, and synonyms for the Taxonomy. The query logs were also one of the sources used to develop use cases for the validation exercises.

Semantic relationships between the Consumer Health Care Taxonomy concepts were made based on trusted sources such as:

- Online symptom checkers from the Mayo Clinic,²³ Cleveland Clinic,²⁴ NHS UK,²⁵ and HealthDirect Australia;²⁶
- the Physician Compare mappings of conditions and symptoms to medical specialties; and
- Google medical search.

Interviews with SMEs and key stakeholders

We started by talking with members of the IDEO team that built the CareFinder prototype and prepared the Design Guide. We particularly focused on the gaps that the IDEO team had identified and documented. Most of these are content gaps in the Medicare.gov datasets, and ideas about how to fill them either by mining the Medicare claims database, or by identifying or creating specific content that consumers often want, such as information about cost and coverage, patient narratives, etc.

The taxonomy team interviewed subject matter experts inside and outside of the government, including focus group sessions with CCSQ and contractor representatives for each of the Compare websites – Dialysis Facility Compare, Home Health Compare, Hospital Compare, Nursing Home Compare, Physician Compare and Supplier Directory. We also talked with groups that do not yet have Compare websites– Inpatient Rehabilitation Facilities/Long Term Acute Care Hospitals and Hospice Care, and with other CMS teams. Our primary interest in all these discussions was to get insights into what consumers want or need to make choices about where to get care.

We looked beyond taxonomy experts to understand how others have tackled consumer-facing websites of various types – particularly UK's National Health Service website - and we looked at SMEs who have great depth and experience – particularly the team from Yale University School of Medicine Center for Outcomes Research & Evaluation – who were able to assist us in navigating subtle challenges of balancing completeness with comprehension.

Finally, we conducted ad hoc interviews with people we encountered in our research to track down background materials, health care professionals such as medical social workers, and friends and loved ones to elicit stories related to the health care decisions they help others make, or that they make themselves.

Interview Groups

- More than 30 CMS staff working on the various Compare websites
- NORC and other Compare data contractors
- Katja Battarbee, co-lead of Phase 1

²² <https://medlineplus.gov/>.

²³ <http://www.mayoclinic.org/symptom-checker/select-symptom/itt-20009075>.

²⁴ <http://my.clevelandclinic.org/health/mysymptomchecker.aspx>.

²⁵ No longer available.

²⁶ <https://www.healthdirect.gov.au/symptom-checker>.

- Frank Funderburk, CMS user research SME
- External health care professionals
- Friends and family who have health care stories

User stories

As we talked to various experts we continued to anchor our development efforts on consumers. Members of the core team created narratives that embodied the types of real-world use cases we have experienced during the course of our lives and past health care work. Our original “user stories” were supplemented by user stories provided by Compare owners, NORC, and others, and provided both the feedstock for developing the taxonomy and a means of validating the efficacy of the taxonomy later on in the project. Here are a couple of typical user stories.

- Sharon is a 52 year old with ESRD who received a kidney transplant 4 years ago. The kidney has recently begun to fail and she needs hemodialysis again 3 times a week. She also recently broke her leg and needs physical therapy 2 times a week. She works as a waitress but is currently unemployed due to her accident, and does not have insurance besides Medicaid. Sharon needs to find a dialysis center close to home since a family member will need to drive her due to her leg injury. She also needs to coordinate her dialysis (3x week) with her physical therapy (2x week).
- Paula is an 85 year-old woman. She was out driving in her neighborhood when all of a sudden she lost her way and she couldn't find her way back home. This episode scared her and her family. She began to worry about her mental capacity, and wondered what kind of specialist she could see who could assess her mental acuity. Paula needs to figure out what type of specialist can help assess her mental acuity, and find a trustworthy specialist who is close to her apartment and covered by her health plan.

The Consumer Health Care Taxonomy

This section describes the purpose, entities and relationships defined in the Consumer Health Care Taxonomy.

Purpose

The Consumer Health Care Taxonomy is focused on supporting the requirements of the CareFinder prototype. The Taxonomy has been designed to function as middleware that translates consumer queries into the language necessary for retrieval of data from Medicare.gov datasets and 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 (i.e., a CareFinder-like 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 2 shows the eleven facets in the Consumer Health Care Taxonomy.

The image shows a screenshot of a web interface displaying the details of a SKOS Concept Scheme. It is divided into two main sections: 'General' and 'Concept Scheme'.

General

- Total concepts:** 2215
- Last Modified:** 07 September 2016 01:33 CET
- Description:** Set of controlled vocabularies (facets) with semantic relationships designed to support the entry points, finding paths of the CareFinder prototype, potential sorting and filtering of search results and tagging original and curated content. Values in CMS datasets are mapped to relevant terms in the taxonomy. (This mapping is done via SKOS relatedMatch, not a custom semantic relationship.)
- Author:** Joseph Busch
- Publisher:** Taxonomy Strategies
- Contributor:** Center for Medicare and Medicaid Services

Concept Scheme

- Body Locations and Systems
- Care Settings
- Conditions
- Content Types
- Demographics and Populations
- Languages
- Medical Supplies and Equipment
- Other Health Topics
- Quality Indicators
- Specialty Areas
- Tests and Treatments

Figure 2-The eleven facets in the Consumer Health Care Taxonomy displayed in the PoolParty Linked Data frontend. In SKOS, a ConceptScheme is an aggregation of concepts or what we call a taxonomy facet. Each of these labels is a TopConcept of the facet.

Entry Terms

The most basic element of a facet or vocabulary are the concepts. One form of the concept is selected as the entry term, or the preferred term²⁷. All other variations are entered as synonyms.

²⁷ In SKOS the entry term is known as the refLabel.

Entry terms have been identified by analyzing search logs for similar sites, related and curated content, popular news sources, and user research including uses cases.

It is important to capture both the technical version of a term and the consumer-friendly or colloquial version(s) of a term. The entry term selected is the one that garners the best results in a GTK content search. Sometimes the best entry term is the technical version, and sometimes it is the consumer-friendly version. One resource used to help determine the version of a term better known to consumers is Google Fight²⁸. This site calculates “visibility” based on the number of results returned by the term or phrase and the number of times the term or phrase is searched for in a month, returning results on a 100 point scale. Figure 2 shows an example of the results of “Spinal tap” versus “Lumbar puncture”.

Post-coordination v. Pre-coordination. The power of a faceted taxonomy is the ability to make different combinations of terms from different facets. This is called post-coordination and creates a type of flexibility that is different from a thesaurus or one simple list. For this reason, we analyze user queries by breaking a multi-word concept into constituent parts. However, as middleware assisting consumers by reflecting their language, many multiple word concepts need to be kept together (that is, pre-coordinated) in this taxonomy. For example, “Hip fracture” is included in the Taxonomy as a pre-coordinated phrase in the Conditions facet.

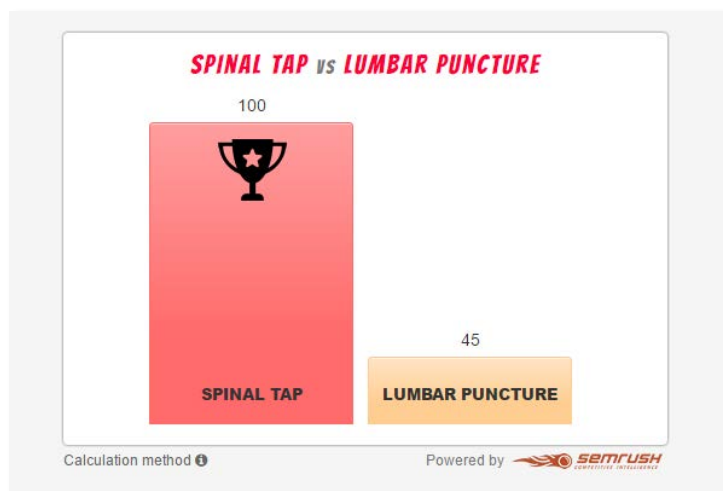


Figure 3-Results of Spinal tap v. Lumbar puncture in Google Fight.

Synonyms and Quasi-synonyms

Synonyms are important for search engines, but they are not usually displayed to the end user. In the Consumer Health Care Taxonomy, variants of the entry term are entered as synonyms.

- True synonyms, e.g. “End-stage kidney disease”.
- Common misspellings.
- Acronyms and abbreviations, e.g., “ESRD”.
- Terms with and without hyphens and special characters, e.g., “End stage renal disease”.
- Irregular plurals.

²⁸ <http://www.googlefight.com/>.

Figure 4 shows synonyms of “End stage renal disease”. In SKOS, a synonym is referred to as an altLabel.

End-stage renal disease
<http://taxonomystrategies.poolparty.biz/CMS3A/Conditions/End-stage-renal-disease>

Alternative Labels

- End stage renal disease
- End-stage kidney disease
- ESRD
- Kidney failure

Broader Concepts

- [Kidney Diseases](#)

Figure 4-Consumer Health Care Taxonomy concept “End-stage renal disease” and its synonyms.

To keep the Taxonomy broad and shallow, children of a term may be entered as synonyms rather than creating another level down. These are called quasi-synonyms. These terms are not true synonyms but are of some value; an additional level down in the Taxonomy is not needed for these quasi-synonyms. Adding quasi-synonyms is also known as posting up or rolling up. Figure 5 shows quasi-synonyms of the Consumer Health Care Taxonomy concept “Canes”. In SKOS, there is no distinction between synonyms and quasi-synonyms. They are all just AltLabels.

Canes
<http://taxonomystrategies.poolparty.biz/CMS3A/Medical-Supplies/Canes>

Alternative Labels

- Offset canes
- Quad canes
- Standard canes

Broader Concepts

- [Canes and crutches](#)

Figure 5-Quasi-synonyms of the concept “Canes”.

Hierarchical Relationships

Hierarchy in the Consumer Health Care Taxonomy is based on the principles of “part of” or “type of”. These are also known as “whole/part” and “is a” relationships. Terms are narrower terms or child terms of a broader term if it represents a part of the broader term or if it is a type or instance of the broader

term. Figure 6 is an example of a “Part of” taxonomy where child terms are parts of the broader concept “Blood glucose monitors and supplies”. In SKOS, hierarchical relationships for a concept are referred to by Narrower and Broader concepts.

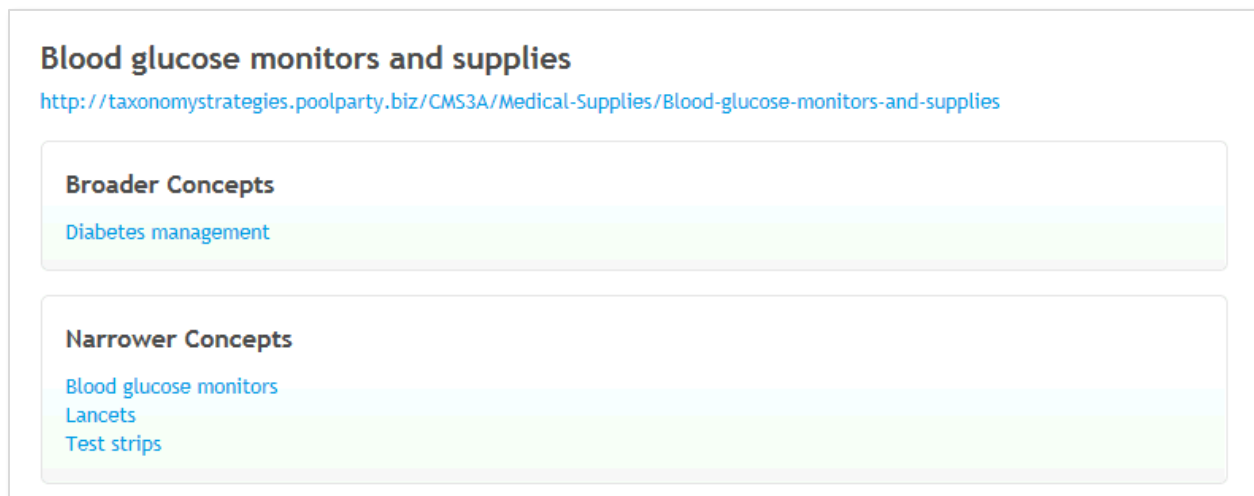


Figure 6-Part of taxonomy for concept "Blood glucose monitors".

Figure 7 is an example of a “Type of” taxonomy where child terms are types of “Fractures”. In SKOS, there is no distinction between the types of hierarchical relationships.

Relationships to entry terms in other facets

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. Currently, six of the eleven facets have one or more relationships to other 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. Table 1 provides an example of the semantic relationships using the facet labels and terms from the facets.

Only relationships that are needed to support consumer health care decision-making have been created.

There is no inheritance of relationships. An entry term from Conditions mapped to a parent in Tests and Treatments is not inherited by its children of the Tests and Treatments entry term. If applicable to both parent and child terms the relationships must be repeated for each.

The semantic relationships are conceived to be in reciprocal pairs, also known as an inverse type associative relation.

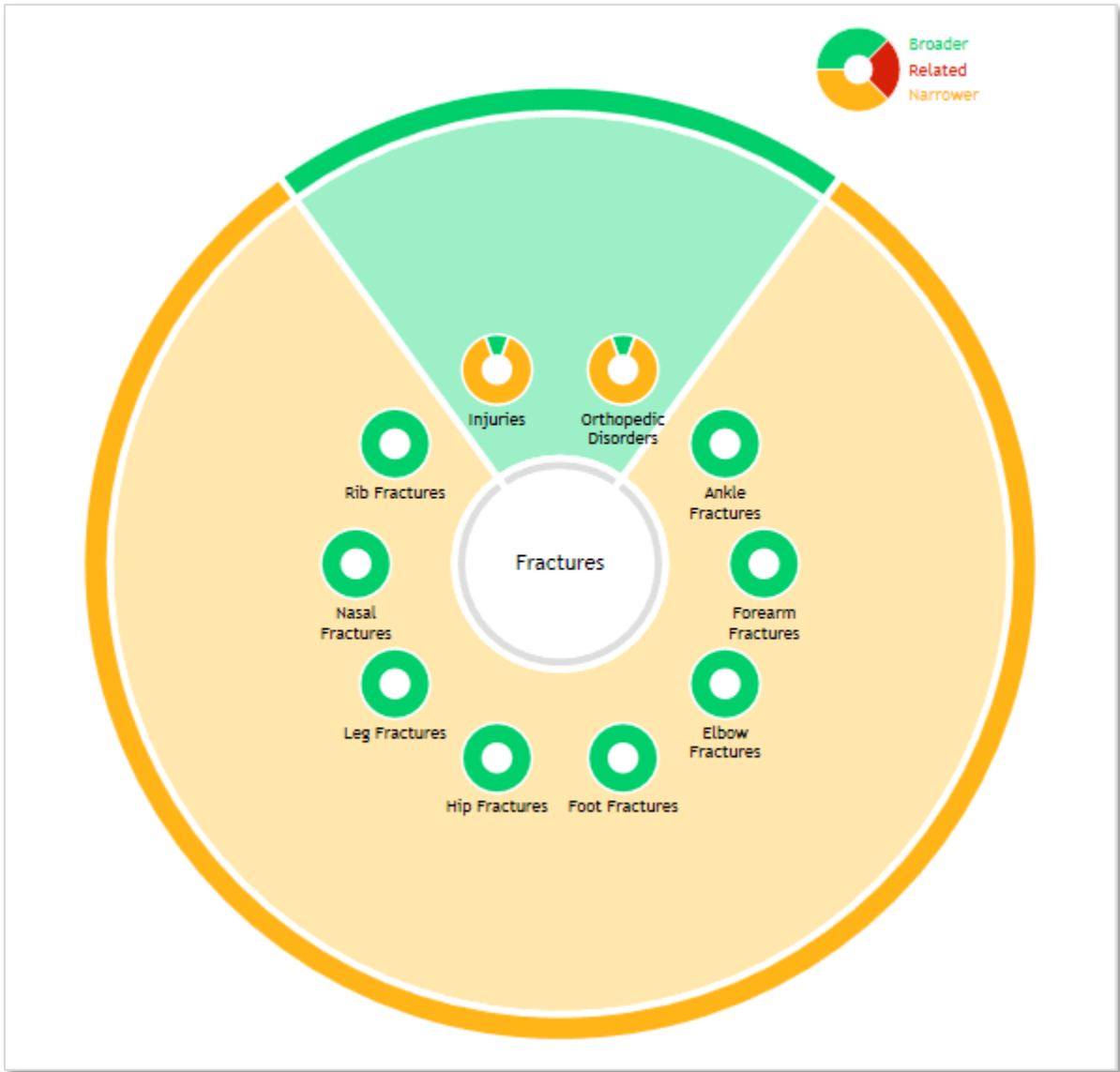


Figure 7-Hierarchical relationships in the PoolParty Linked Data frontend visual interface.

Semantic Relation			Inverse Semantic Relation		
Facet Class		Facet Class	Facet Class		Facet Class
Body Locations and Systems	is affected by	Conditions	Conditions	affects body location	Body Locations and Systems
<i>Kidneys</i>	<i>is affected by</i>	<i>End-stage renal disease</i>	<i>End-stage renal disease</i>	<i>affects body location</i>	<i>Kidneys</i>
Conditions	has treatment of	Tests and Treatments	Tests and Treatments	is treatment for	Conditions
<i>End-stage renal disease</i>	<i>has treatment of</i>	<i>Dialysis</i>	<i>Dialysis</i>	<i>is treatment for</i>	<i>End-stage renal disease</i>
Conditions	is concern of	Specialty Areas	Specialty Areas	is concerned about	Conditions
<i>End-stage renal disease</i>	<i>is concern of</i>	<i>Nephrology</i>	<i>Nephrology</i>	<i>is concerned about</i>	<i>End-stage renal disease</i>
Conditions	needs medical supply	Medical Supplies and Equipment	Medical Supplies and Equipment	is needed for condition	Conditions
<i>End-stage renal disease</i>	<i>needs medical supply</i>	<i>Dialysis Equip. & Supplies</i>	<i>Dialysis Equip. & Supplies</i>	<i>is needed for condition</i>	<i>End-stage renal disease</i>
Care Settings	is location of treatment	Tests and Treatments	Tests and Treatments	is treatment provided in	Care Settings
<i>Dialysis Facilities</i>	<i>is location of treatment</i>	<i>Dialysis</i>	<i>Dialysis</i>	<i>is treatment provided in</i>	<i>Dialysis Facilities</i>
Care Settings	specializes in	Specialty Areas	Specialty Areas	is specialty of	Care Settings
<i>Dialysis Facilities</i>	<i>specializes in</i>	<i>Dialysis Services</i>	<i>Dialysis Services</i>	<i>is specialty of</i>	<i>Dialysis Facilities</i>
Medical Supplies and Equipment	is used in treatment	Tests and Treatments	Tests and Treatments	uses medical supply	Medical Supplies and Equipment
<i>Dialysis Equip. & Supplies</i>	<i>is used in treatment</i>	<i>Dialysis</i>	<i>Dialysis</i>	<i>uses medical supply</i>	<i>Dialysis Equip. & Supplies</i>
Specialty Areas	includes treatment of	Tests and Treatments	Tests and Treatments	is part of practice area	Specialty Areas
<i>Nephrology</i>	<i>includes treatment of</i>	<i>Dialysis</i>	<i>Dialysis</i>	<i>is part of practice area</i>	<i>Nephrology</i>
Care Settings	has focus of condition	Conditions	Conditions	is focused on in setting	Care Settings
<i>Dialysis Facilities</i>	<i>has focus of condition</i>	<i>End-stage renal disease</i>	<i>End-stage renal disease</i>	<i>is focused on in setting</i>	<i>Dialysis Facilities</i>
Body Locations and Systems	location is treated by	Tests and Treatments	Tests and Treatments	treats body location	Body Locations and Systems
<i>Kidneys</i>	<i>location is treated by</i>	<i>Dialysis</i>	<i>Dialysis</i>	<i>treats body location</i>	<i>Kidneys</i>

Table 1-Consumer Health Care Taxonomy semantic relationships and examples.

Mapping to Medicare.gov dataset values

In order to show relevant information on types of care providers and settings related to a consumer's keyword query or other chosen path, relevant taxonomy terms have been mapped to relevant terms from the Medicare.gov datasets managed by CMS. Key to these mappings is a balance between enough to be useful to the consumer but not so many that the consumer is overwhelmed with too many results or results that do not make sense to the consumer. The relevant terms from the Medicare.gov datasets are identified and made available separately from the actual datasets due to the heterogeneous nature of the current datasets and other complicated data concerns.²⁹ In SKOS, mapping properties are used to align different concept schemes. We use the relatedMatch property to state an associative mapping between Consumer Health Taxonomy concepts and Medicare.gov dataset values.

Figure 8 shows an example where there is a specialty from the Physician Compare dataset (*Speech Language Pathologist*) mapped to the similar concept (*Speech Pathology*) in the Specialty Areas facet. *Speech Pathology Services* is a specialty from the Home Health Compare dataset that is also mapped to *Speech Pathology*. In developing the Taxonomy, initial mappings to values in the Medicare.gov datasets were made based on ease of identification by non-medical experts. Thus, the initial mappings are narrow in scope. These mappings need to be reviewed and extended with the guidance of subject matter experts.

The screenshot displays a taxonomy entry for 'Speech Pathology'. At the top, the title 'Speech Pathology' is followed by a URL: <http://taxonomystrategies.poolparty.biz/CMS3A/Specialty-Areas/Speech-Pathologist>. Below this, there are three sections: 'Alternative Labels' listing 'Speech and Language Therapist', 'Speech Language Pathologist', 'Speech Pathologist', 'Speech Therapist', and 'Speech Therapy Specialist'; 'Broader Concepts' listing 'Speech, Language and Hearing Services'; and 'Related Matching Concepts' listing two URLs: http://taxonomystrategies.poolparty.biz/CMSDatasetValues/Compare-Dataset/Speech-Language-Pathologist_2 and http://taxonomystrategies.poolparty.biz/CMSDatasetValues/Compare-Dataset/Speech-Pathology-Services_2.

Figure 8—Example of mapping a specialty—Speech Pathology—to Medicare.gov dataset values.

Observations from the Consumer Health Care Taxonomy Project

This section includes some observations from the Consumer Health Care Taxonomy project.

²⁹ For the initial build a large subset of values from the [Medicare.gov](http://www.medicare.gov) datasets were identified and held separate from the Consumer Health Care Taxonomy in the taxonomy management tool, PoolParty.

- Consumer healthcare related decision-making behavior is different from clinicians.
 - Focus on the problem to be solved: Translate consumer queries into the language necessary for retrieval of data from Medicare.gov datasets and Good to Know web content.
 - Exhaustivity is not a requirement.
- While there are many healthcare-related technical KOS available, consumer-friendly terminology is generally not available from authoritative sources.
 - A lot of work is required to compile a useful KOS from many sources.
 - Documentation of editorial guidelines supports this activity and helps to make it scalable.
- A small set of extensible taxonomies and custom semantic relationships are sufficient to develop the domain model.
 - A concise set of subject predicate object relationships, e.g., Condition is_concern_of Specialty Area.
 - SKOS is not intended for encoding more complex ontologies beyond thesaurus relationships (hierarchy, equivalent and generic associative), so a custom schema was developed for specific associative relationships.
- A strategy to setup separate concept schemes for the Consumer Health Care Taxonomy and the Medicare.gov datasets controlled vocabularies provided flexibility and extensibility.
 - SKOS relatedMatch was used to map across the concept schemes.
 - KOS management tools are immature in their capacity to accurately and efficiently batch import and export KOS, interim taxonomies and semantic relationships.