**Some Taxonomy Definitions**

**Taxonomy:** A collection of terms and relationships between terms used to describe a domain (and hopefully created to address a real business problem). The terms or relationships may or may not have rich attributes associated with them.

There are standards that have been developed. You can find some excellent information in the ANSI/NISO Z39.19

<http://www.niso.org/apps/group_public/download.php/12591/z39-19-2005r2010.pdf>

The work here reflects this but is not completely bound by the standard.

Some common types of collections:

|  |  |  |
| --- | --- | --- |
| Structure  | Description | Example |
| Controlled value lists | A flat list of terms with no explicit relationships. | Drop down list on web form. |
| Taxonomy | A collection of terms with hierarchical relationships between them. All terms are descendents of a common ancestor. Terms may have more than one parent. Terms may have properties like scope notes | Product hierarchy, Geographic hierarchy, etc. |
| Thesauri | A taxonomy with Entry (synonym) and Associated (see also) relationships. The relationships may be typed (named).  | Often used in internal systems and may be used with search systems or to power a corporate glossary. |
| Ontologies | A collection of concepts, relationships, attributes, and rules describing a domain. These are used to model complex domains | Often used in healthcare and life sciences to support sophisticated inferences and conceptual models |
| Glossary | A list of terms, generally in alphabetical order, with descriptions, synonyms, and other potential metadata. | Dictionary, Source for Tool tips, etc. |
| Folksonomies (according to some) | User generated lists of words, generally collected via uncontrolled tagging activities | User tags |

**Parts of a Taxonomy**

**Terms:** The core element of a taxonomy, all of the concepts that are being described!

**Term Properties:** Terms can have descriptive properties. Common properties are listed below, but the list can be modified depending on the requirements.

* Label: a word or phrase used to present this concept to end users
* System name: generally a human readable version of the term that is unique in the system
* Scope Notes: a definition of the concept that is valid within the context of this domain
* IDs: unique identifiers used by systems. Generally there is only one ID per system, but if the term is used in multiple systems, then it can be necessary to track all those IDs in one place
* Dates: time stamps describing predefined actions
* Others, such as:...

Example for the term “Soup Spoon”:

* Label: Soup Spoon
* System Name: soup-spoon
* Scope Note: an implement consisting of a small, shallow oval or round bowl on a long handle, used for eating liquids like soup.
* System ID: 893274h7d7
* External ID: spoon\_456
* Creation Date: 4/14/98
* Last Modified Date: 9/9/2015
* Last Published Date: 3/7/2017
* Tagging Notes: Use this for spoons that are identified by the manufacturer as soup spoons, or for spoons that are in a set where they are larger spoon.
* Use in Sets: True

Sample types of terms:

* Preferred term - generally used for tagging
* Entry Term - synonyms, abbreviations, acronyms. These are directly associated with a preferred term
* Leaf Node - a term with no children.
* Placeholder Term - often used to provide an important structural component of the taxonomy, even though it would not be used for tagging or other purposes.

**Relationships:** Therelationships between concepts may have properties as well.

Hierarchical: Broader and Narrower nodes – a “tree” structure

* Parent/Child - Stereo System → Speakers
* Part/Whole - North America → Canada
* Genus/Species - Dog → poodle

Entry or Equivalence: “also known as” relationships

* Synonyms - Automobile → Car
* Antonyms - Automobile →
* Abbreviations - Automobile → Auto
* Common Misspellings - Automobile → Automoble, Autombile...

Associative – “see also” relationships used to move horizontally in the taxonomy



**Rules:** a way to encode some logic into your taxonomy and control what gets used and when.

* Examples of rules applied to properties:
	+ “All terms must have a label.”
	+ “All terms must have scope note.”
	+ “All term labels must be capitalized”
* Examples of rules applied to relationships:
	+ “The term on the other end of a ‘Has Product’ relationship must be a product.”
	+ “The entity type ‘Product’ may not have a child.”
	+ “The entity type ‘Recipe’ may have children.”