

B E Y O N D

Think Beyond. Think Mphasis.

WHITE PAPER



## Taxonomy Harmonization – A Tool For ‘Improved Findability’

Brijesh R Marathe,

Technical Architect in ‘Enterprise Search Practice’

August, 2010

## Table of Contents

1	Need for Taxonomy	2
2	Categorization and Classification	2
3	Building Taxonomy	2
4	Categories of Taxonomies	3
	User-based Taxonomies	3
	Market-based Taxonomies	3
5	Taxonomy Harmonization	3
6	Localized Taxonomy	4
7	Conclusion	4
8	Suggested Reading	4

## 1 Need for Taxonomy

In today’s virtual and online world most of the organizations are in a pursuit to promote their businesses, services and products by maintaining websites for their brands. It is essential to promote these businesses, services and products through increased traffic to these websites. Major Search Engines act as gateway to increase the traffic to your websites that are promoting your services and products.

The advent of Search Engine Marketing or Internet Marketing has helped the business to understand, and therefore focus, on marketing and advertising through web search engines. Solitary objective is to be found on the web by thousands of users, your potential customers, searching for services and products which you provide and want to promote online.

In today’s highly competitive markets, search engines are playing a major role as a marketing medium hence the ‘ability to be found’ on search engines’ result page (SERP), is not just a nice to have feature but it drastically affects the top line<sup>1</sup> of your business. Along with search engine marketing (SEM), search engine optimization (SEO) has become indispensable and a de-facto to build web sites to increase the findability. How, then, can it be improved?

## 2 Categorization and Classification

How do users search the internet using search engines? They use **keywords**, free text that describes what they are searching for. Search engines show the most relevant web pages for these keywords in their result pages. Generally most users look at a couple of initial result pages for what they are searching (services and products). If they don’t find what they want (i.e. not satisfied), they either redefine their search or they abandon the search and move on. These keywords, therefore, are vital for you to get noticed by these potential customers i.e. for the ‘ability to be found’.

Search Engine Optimization (SEO) is a process of optimizing web sites to improve the traffic from search engines in organic way. SEO is a harmonization of the way search engines work and what people search for. That is why the keywords from taxonomies along with effective search engine optimization techniques can help to increase the traffic to your site. SEO leverages the keywords and enforces to put them in specific and strategic locations on the web pages where search engine crawlers (bots) look while extracting the web page content. For a successful website, a highly important aspect is to align the content of the website (each page of the website) with the ‘Keywords’ that customers might use for and land on your site where you expect or want them to land. Site content classification and categorization lets you align the content of your website in the most appropriate and intuitive manner.

Taxonomy plays a significant role in classification and categorization of the site content. Taxonomy is the hierarchical organization of a particular set of information for a particular

purpose. In the context of your web presence, taxonomy is a classification system that can guide how the web pages are organized on a website. It can help customers quickly search the site and locate the items (services and products) they’d like to purchase easily. With taxonomy in place, when users perform search, they quickly find the exact information they need.

## 3 Building Taxonomy

Your site visitors use search engines to discover the websites for the items that they are searching. Upon landing at the correct web site, they may spend time to locate related items on the web site by navigating through the website. A successful taxonomy should, for that reason, be planned and designed both for **searching** and **natural browsing**.

Both these aspects have to be taken into consideration to build an effective and successful taxonomy. Building taxonomy to organize the web site content is as much an art as it is a science. One should take an Information Architectural approach to design and build the Taxonomy. There are substantial differences between how search engine crawlers (bots) consume the site content and how people consume it. Information Architecture outlines a method to categorize the information into a coherent structure that most people can understand quickly. Also, Along with the search engine optimization (SEO) techniques, Information Architecture makes search engine crawlers (bots) to consume the site content in most efficient way hence results in drastically improved findability.

Always keep in mind that taxonomies are not inert; they need periodic review, they need to be in pace with time, your current market and user behavior trends. Building taxonomies is an expensive and time consuming process. Therefore, the way taxonomies are planned, designed, built to allow for reviews and adjustments becomes very important.

With effectively built taxonomies, your site visitors will be empowered with an influential ‘Find-ability’ experience, enabling them to discover information they are looking for in a natural and consistent manner.

---

<sup>1</sup> ‘Improved Findability Leads to Positive Top Line Impact’ - MphasiS white paper by Chirag Gandhi

## 4 Categories of Taxonomies

The taxonomies generally and most commonly planned, designed and built are the Concept based taxonomies. This is a traditional line of attack to build taxonomy. These taxonomies exploit all possible terms of the concept for which taxonomy is built. For example ‘Organic Beauty Products’ could be a concept for which the taxonomy can be built to classify and categorize the site content. To build these taxonomies you will need a Subject Matter Expert, and it is a manual process. Research Analysts, your Marketing team and communication team can also contribute to build concept based Taxonomy. In addition, content held in a content management system can contribute to your taxonomies through the meta-tagging capabilities of content management systems. Based on our experience, we think that apart from this traditional category of taxonomies there are even more important and generally overlooked categories of taxonomies and they are:

- User Based Taxonomy
- Market Based Taxonomy

### User-based Taxonomies

Whenever visitors search for services or products on your site, they are indicating what is important to them; they are trying to tell you what terms are significant to them. Actually they are indicating what products they want to buy; they are pointing to let you know what they are expecting from your website. If you are not paying attention to these then, unfortunately, you are losing out your **User-based Taxonomy**, also termed as user behavior based taxonomy. Investigating and analyzing internal search engine logs, which is known as Site Search Analysis or Log Analysis will lead you to discover the user based taxonomy. The set of the top queries or search keywords from your internal site search engine is what, we call User based taxonomy; the most important taxonomy you should adopt to. Google, the giant database of user search keywords, uses a tool to select the keywords for a concept based on its algorithms applied on a huge amount of user search keywords through users’ action logs. These keywords are also strong contributors to the user based taxonomies because with these keywords users are searching for the products or services that are associated with specific concepts. For example “Health Insurance” is a concept for which Google can provide the keywords and terms that users searching across the globe from its various logs.

### Market-based Taxonomies

It is crucial to realize the taxonomy trends in the industry to use the terminology and keywords appropriately in your taxonomy. The understanding of how your competitors are using taxonomy is equally important in improving the content of your site and choosing competitive terminology. This taxonomy is what we call a Market Based Taxonomy. It is based on the trends in the market and from your competitors so Marketing Teams and Research analysts should be called upon to assist in building such

taxonomies. Exploring and evaluating the competitors’ websites using automated tools allows you to generate a taxonomy used by your competitors which will help you create a Market Based Taxonomy.

## 5 Taxonomy Harmonization

All these discussed categories of taxonomies are important and need to be considered to design and build an effective taxonomy. The art of aligning your taxonomy with market and user based taxonomies is called Taxonomy Harmonization. Each of the taxonomies can be represented as a set of keywords and or concepts. When all the taxonomy sets are superimposed we get a visual representation as shown in figure 1. To achieve a harmonized taxonomy we try to increase the intersection between all the sets as represented by the areas marked with ‘1’ in the figure 1.

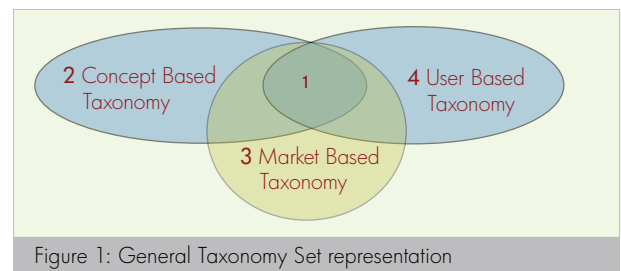


Figure 1: General Taxonomy Set representation

1	The well-adjusted intersection is the harmonized state of all categories of taxonomies which gives Harmonized Taxonomy.
2	Terms which are part of your taxonomy but they are not appearing in market taxonomy and user taxonomy. For these terms you need to be careful during harmonization because these terms indicates specialization in your taxonomy
3	Terms present in the market taxonomy (used by your competitors’), but not being used by you. It is always beneficial to incorporate these terms in your taxonomy.
4	Terms being used by the users (during internal site search), but not present in your taxonomy. These terms represent opportunities to align yourself with your customers better.

Achieving the harmonious state between all these categories of taxonomy is really an art and can be effectively achieved only if there is a good understanding of your business, the market and the users.

Your taxonomy should also always be considered “under construction” and the search experience should be monitored to improve the taxonomy at regular intervals.

## 6 Localized Taxonomy

Localization of the taxonomies is a very significant element of Taxonomy Harmonization. You need to pay proper attention to Country and language (Locale) while harmonizing the taxonomies.

The localization of taxonomy is the adaption of keywords that are popular in a particular region where products or the services are intended to be sold for the business to prosper. For example, for a specific brand to be sold in US region you will have to select the keywords in accordance with the US region but you will select different keywords for the UK region based on the customs, culture, sentiments, general market trends etc.

You will have to consider the language aspect as well while building taxonomy. The country where two different languages are used more commonly, based on density of the population, you will have to build the taxonomy for both the languages. As per our experience, you should harmonize the taxonomies for a Brand in accordance with the Country and the language as shown in figure 2.

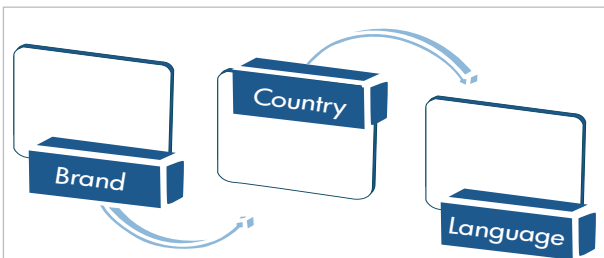


Figure 2: Localization of Taxonomy

## 7 Conclusion

User Based Taxonomies and Market Based taxonomies are the companion of the traditional concept based taxonomies, therefore, they should go hand in hand to outline a Harmonious state. Taxonomy Harmonization facilitates to achieve this and ensures that all types of users' needs will be considered in the construction of the taxonomy. The accurate classification of the Site Content in accordance with the Harmonized Taxonomy improves the ability to find information on your web site. The 'ability to be found' improves your top line accordingly.

At MphasiS, the Enterprise Search Practice team provides innovative solutions to customers through a combination of technology know-how, domain expertise and visual tools to manage the entire process of taxonomy harmonization.

## 8 Suggested Reading

1. Building Enterprise Taxonomies, Darin L. Stewart, Mokita Press, 2008.  
ISBN: 978-1419693265
2. The Most Important Taxonomy for your Web Site, New Idea Engineering, on the web at:  
[http://www.enterprisearchblog.com/2007/04/the\\_most\\_import.html](http://www.enterprisearchblog.com/2007/04/the_most_import.html)
3. Taxonomy Watch, on the web at  
<http://taxonomy2watch.blogspot.com/>
4. Business Metadata and Enterprise Search, The Data Administration Newsletter, on the web at  
<http://www.tdan.com/view-articles/4579>
5. Better Living Through Taxonomies, Digital Web Magazine, on the web at  
[http://www.digital-web.com/articles/better\\_living\\_through\\_taxonomies/](http://www.digital-web.com/articles/better_living_through_taxonomies/)
6. Taxonomy Community of Practice Wikispace  
<http://taxocop.wikispaces.com/>

## Contact us

### USA

MphasiS  
460 Park Avenue South  
Suite # 1101, New York  
NY 10016, U.S.A.  
Tel: +1 212 686 6655  
Fax: +1 212 686 2422

### UK

MphasiS  
88 Wood Street  
London EC2V 7RS, UK  
Tel: +44 208 528 1000  
Fax: +44 208 528 1001

### AUSTRALIA

MphasiS  
410 Concord Road  
Rhodes, NSW 2138, Australia  
Tel: +61 290 221 146  
Fax: +61 290 221 134

### INDIA

MphasiS  
Bagmane Technology Park  
Byrasandra  
C.V. Raman Nagar  
Bangalore 560 093, India  
Tel: +91 80 4042 6000  
Fax: +91 80 2534 6760

## About MphasiS

MphasiS is a global service provider with \$1B in revenues, delivering technology based solutions to clients across the world. With currently over 39,000 people, MphasiS services over 200+ clients in Banking and Capital Markets, Insurance, Manufacturing, Communications, Media & Entertainment, Healthcare & Life Sciences, Transportation & Logistics, Retail & Consumer Packaged Goods, Energy & Utilities, and Governments around the world. Our competency lies in our ability to offer integrated service offerings in Applications, Infrastructure Services & Business Process Outsourcing capabilities. We are uniquely positioned to service our clients with best cost-performance. To know more about MphasiS, log on to [www.mphasis.com](http://www.mphasis.com)

MphasiS and the MphasiS logo are registered trademarks of MphasiS Corporation. All other brand or product names are trademarks or registered marks of their respective owners. MphasiS is an equal opportunity employer and values the diversity of its people. Copyright © MphasiS Corporation. All rights reserved.