



Københavns Universitet

Company Taxonomy development

Lund, Haakon; Ørnager, Susanne

Published in:
Aslib Journal of Information Management

DOI:
[10.1108/AJIM-09-2015-0140](https://doi.org/10.1108/AJIM-09-2015-0140)

Publication date:
2016

Document Version
Peer reviewed version

Citation for published version (APA):
Lund, H., & Ørnager, S. (2016). Company Taxonomy development: The case of an international emergency response organization. DOI: 10.1108/AJIM-09-2015-0140

Company Taxonomy development : the case of an international emergency response organization

Purpose

– This study aims to explore theoretically and empirically the understanding and implementation of an information taxonomy in the UN organization World Food Programme (WFP) by analyzing users' information behavior and by establishing a minimum set of cross silo metadata (taxonomy).

Design/methodology/approach

- The study implies the use of both qualitative and quantitative methods. This includes desk review of key documents and interviews with information architecture staff from various WFP units; data collection carried out as semi structured staff interviews in WFP; log analyses of search log files from WFP intranet portal (WFPgo) from September – November 2013, the results were applied and a suggested taxonomy tested at workshops conducted for the staff in HQ.

Findings

– The results reveal an organization with a high demand for easier access to information and knowledge, greater internal collaborations and stronger links with various sources of knowledge. Staff participating in the various workshops pointed out that work processes as well as the human resources component cannot be left out of a solution development.

Originality/value

– There has been little research carried out on current taxonomy projects in corporate environments and international emergency response organizations and few has touched on how knowledge organization systems can enhance or constrain staff's ability to access online content.

Keywords

- Information and knowledge management, Company Taxonomy, Knowledge organization systems, Log analyses, Card Sorting, Information architecture, Emergency response organization, United Nation's (UN) World Food Programme (WFP).

Paper type

- Research paper

1. Introduction

Most large organizations use intranet for internal communication and knowledge exchange. Regrettably, a companywide policy for defining the scope of the implemented intranet and in several cases a suitable communication strategy for employing information on the intranet is missing (White, 2010). This often results in reluctance by staff to use the intranet actively for seeking information since the experiences are frequently unproductive and time-consuming. This may lead to solutions only supporting the information needs within isolated parts or departments of the organisation in what can be described as siloed information and systems i.e. information systems created by one organizational unit in isolation, with the intended use by this unit (Fenwick et al., 2009; Tariq et al., 2014).

Poor information architecture of the main corporate intranet portal is another barrier for user acceptance. The information architecture is here regarded as the structure of the intranet including use of metadata, navigational structure, search facilities, organisation of document repositories/management system and page layout. If users encounter poor information architecture they may apply for alternative information access strategies like the use of personal contacts, the use of external search engines (e.g. Google) and the use of external information systems such as Facebook, Twitter etc.

The Technology Acceptance Model (TAM) (Davis, 1989) has become widely used as a model for explaining users acceptance of information systems. The TAM model operates with two core concepts, 'perceived ease of use' and 'perceived usefulness' as explanations to what degree a system is accepted. The 'perceived ease of use' is the amount of effort a user have to invest in using a system and systems 'perceived usefulness' is to what degree a user think he will benefit from using the system. If users finds that a system fails on either the perceived ease of use or perceived usefulness (or both) it can lead to users rejecting the system.

According to Stocker et al. (2014) only a few research papers have been published on how employees search for information within an organization including how useful users finds intranet search facilities or enterprise search (ibid., p. 2). The company Findwise (2014) reported in a survey with 277 participants that 43% of the respondents found it hard to very hard to find information within the organisation. Of these 50% of the respondents were from large organisation (more that 1000 employees). When asked what the obstacles users encountered app. 40% indicate inconsistency in tagging content and app. 45% lack of adequate tags i.e. metadata.

Taxonomy is a conceptual framework for structuring content so organizations and employees can easily locate what they need. When used in the framework of the internet, taxonomy refers to the effective structuring of content within a defined scope that facilitate easy and accurate access and it is understood as a component of the information architecture (ANSI/NISO Z39.19, 2005). A

corporate taxonomy is an important tool for establishing an efficient information architecture (Gilchrist, 2001; Gilchrist, 2003; Chaudhry and Jiun, 2005; Broughton, 2006a).

However, there is no single agreed definition of a taxonomy either within information architecture (Morville and Rosenfeld, 2007) or information science. It may be understood in a number of ways as theories or frameworks. According to Garshol (2004) the term taxonomy has been widely used and abused to the point that when something is referred to as taxonomy it can be just about anything, though usually it will mean some sort of abstract structure. He defines taxonomy to mean a subject-based classification that arranges the terms in a controlled vocabulary into a hierarchy without doing anything further.

Lambe (2007) claims that common definitions of taxonomies often limit themselves to hierarchies. He states that in its basic definition, a taxonomy is a structured set of names and descriptions used to organize information and documents in a consistent way.

Whittaker and Breininger (2008) define a taxonomy as a controlled vocabulary, in which each term usually holds hierarchical relationships, which means that a taxonomy imposes a topical structure on information.

The term taxonomy is used by Hedden (2010) as a framework. In the narrow sense, to mean a hierarchical classification or categorization system, and in the broad sense, in reference to any means of organizing concepts of knowledge.

A taxonomy is according to the Special Interest Group of the American Society for Indexing (Taxonomies, 2013) defined as a controlled vocabulary with a hierarchical structure, with the shared understanding that there are different definitions of a hierarchy. Terms within a taxonomy have relations to other terms within the taxonomy. The term taxonomy tends to be used to refer to two different things:

- a tree-hierarchical controlled vocabulary lacking more complex relationships found in thesauri or ontologies, or
- any kind of controlled vocabulary, especially when applied to the world of enterprise content management and Website information architecture, rather than library science literature retrieval.

The researchers Garshol (2004) and Whittaker and Breininger (2008) agree that a taxonomy is a controlled vocabulary with hierarchical relationship while Lambe states that this limits the definition. Hedden and the Special Interest Group of the American Society for Indexing agree with the first scholars, however, they provide also a second understanding of a taxonomy which relates to Lambe's claim of limitation.

Hedden's broad framework, that a taxonomy refers to any means of organizing concepts of knowledge (Hedden, 2010) is supported by Pincher (2010a) and Lambe (2011). The first argues that, without a taxonomy designed for storage and management or one that supports better searching, all types of management systems in an organization are nearly useless. The latter clearly mentions that

knowledge taxonomy can provide standard and common understanding of subjects in organizations.

In an organization where shared folders are a bit messy it can be hard to locate information. It usually takes less time to ask the person responsible about what cannot be found. Here, the taxonomy should help to negotiate inconsistencies and harmonize vocabularies. A common taxonomy helps an organization understand and get access to the information it holds as well as missing information. It also allows the use of related information that was previously divided into separate areas of management.

“Ultimately, taxonomies need to reflect the working environment and culture of the organization for which they are created. Since working environments change continuously, taxonomies should also be flexible and adapt to the changing environment” as Pelini and Jones (2011a, p. 15) put it.

2. Paper objectives

This paper reports on some of the main findings from a comprehensive study on the internal information architecture of United Nation’s (UN) World Food Programme (WFP). The study took place in a 12-month period beginning October 2013 and was conducted by the authors of this paper and involving a number of students from University of Copenhagen, Royal School of Library and Information Science.

The WFP is the largest UN organization with app. 14.000 employees worldwide. The majority of staff is field based either in regional or local offices and about 1200 persons are employed at WFP headquarters (HQ) in Rome, Italy. The scope of the organization is to deliver emergency food aid to victims of natural disasters, war and civil conflicts but also to provide food when an emergency situation has occurred and afterwards to help communities rebuilding the infrastructure. In 2013 WFP delivered aid to approximately 81 mill people (World Food Programme, 2014). The WFP emphasis on emergencies does create a number of information management challenges in the organization since the main focus is on first response and action within hours of a new crisis and not specifically on information management and access. In this framework, WFP can be characterized as an emergency organization where an emergency is considered a project with its own life cycle.

Because of the conditions outlined above, a growing frustration has been observed among the staff relating to difficulties in accessing information when needed and the wish among management for an improved knowledge sharing and to prevent the problems with various offices developing siloed information systems inaccessible for other users (Fenwick et al., 2009; Tariq et al., 2014; Wang and Noe, 2010). The objective of this case study was therefore to evaluate the existing information architecture and if necessary, to propose solutions for improving the current information architecture by implementing a taxonomy. Therefor the overall research question is:

- RQ To consider the information architecture for WFP on a superior organizational level not only to develop the organizational system and structure for an intranet but to establish a corporate and common terminology for this global organization.

To operationalize the research question two survey questions were identified which also provide guidance for choosing the data collection methods:

- RQ1 To analyse users' (WFP staff and volunteers) search for and access to information on intranet and internet.
- RQ2 To analyse WFP's information architecture and provide recommendations for revised structures which can help users finding the information they require. This involves organization and structure of information, navigational structure and a categorization system – a taxonomy.

A number of methods were applied for analysing the information system and uncover a somewhat inconsistent information architecture and subsequently to construct a taxonomy (Conway and Sligar, 2002). The methods are well known and tested in the information science field. The study is divided into more phases which follows the research questions 1 and 2. Each phase has different data collection methods attached.

RQ1 started with an IT feasibility study of the intranet in WFP followed by data collected from semi-structured interviews (Kvale and Brinckmann, p. 135, 2009) among WFP staff in the field and HQs about the different understandings in utilizing the intranet. A quantitative log-file analysis of WFPgo's (WFP's intranet) search-log from September to December 2013 ended the first phase.

RQ2 involved two workshops on taxonomies for the staff in HQs. The first workshop with 27 participants was a closed card-sorting experiment in combination with observations of the participants (Morville and Rosenfeld, 2007). The second workshop (27 participants) was on development of synonym rings.

3. Feasability study of intranet use in WFP (RQ1)

The current information architecture in WFP includes a number of different information systems with an intranet portal as the main entry point. The intranet portal called WFPgo is based on a Web Content Management System and contains a large quantity of Web pages. The default home page of WFPgo (see figure 1) is organised with a top menu bar containing 8 drop down menus for site wide navigation. One menu is a Topic -menu with access to topic sites for different units within WFP who is also authoring the content. Next is a Go Guide – menu organised by subjects and linking to pages authored by the

communication branch. On the left side is a list of quick links to selected pages and information systems is found. Pages available as quick links are also found as menu items at the GoGuide -menu. Finally an A-Z site index is available at the left side menu.

Further the WFPgo are indexed and searchable. The portal does also provide access to a number of other information systems including a corporate wide document management system, individual siloed information systems within departments and a number of Wikies . Content on the WFPgo portal is searchable through an enterprise search engine (Google Appliance Box) and only selected documents from the document management system are indexed by the WFPgo search facility. The document management system does, however, provide its own search function.

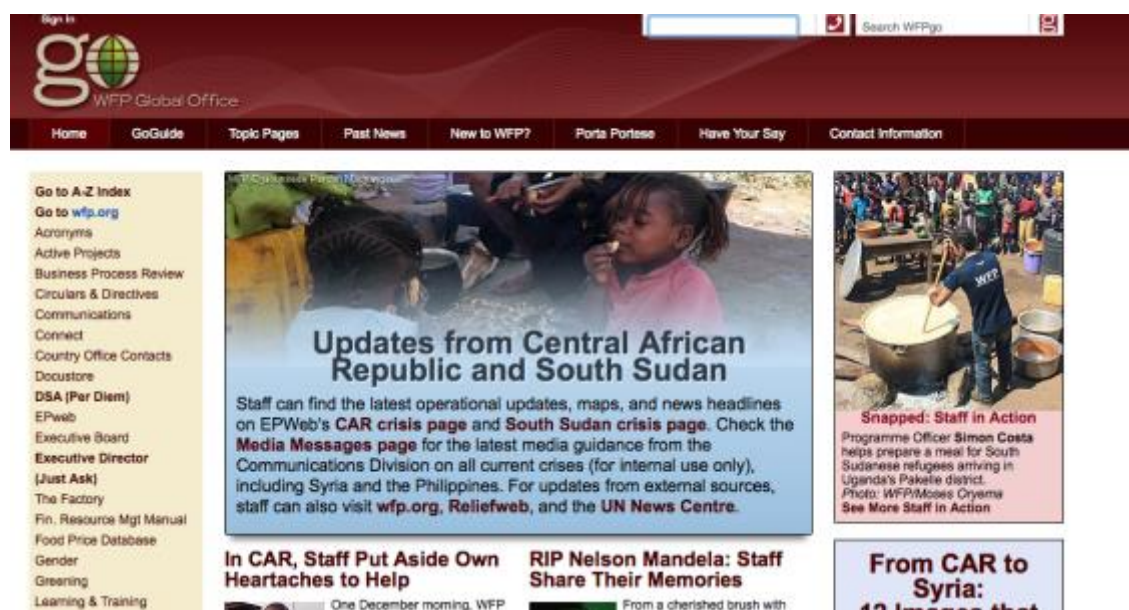


Figure 1 : WfpGo homepage

4. Data collection (RQ1)

To identify issues with the current information architecture within WFP and subsequently to support the taxonomy development both qualitatively and quantitatively data collection methods were applied. The scope of the study was twofold:

- to identify navigational issues in the existing information architecture

And based on these findings

- to provide a platform for the development of a taxonomy to facilitate an improved access to documents

During the first half of 2013 a number of internal inquiries have been conducted by Boston Consulting Group (BCG) to uncover user frustration with the information architecture. The main findings related to difficulties in locating relevant information within WFP system when using the exiting information systems and the need for retrieving a topic expert. The latter is imperative because of the frequent rotation of staff between HQ and field stations. Difficulties in finding relevant (or the right) information are considered the main issues among staff.

To endorse the above assumption we organized interviews with ten (10) participants from HQ and field offices. The interviews were conducted as semi-structured interview and an interview guide was used (Kvale and Brinckmann, 2009). Nine (9) interviews took place in HQ and one (1) was a telephone interview with a respondent from a field office. The analysis of the interviews did confirm that users have difficulties in retrieving relevant information and that this is due both to navigational (menu structure) issues and poor search engine performance.

Respondent 2 on navigational issues:

'The structure is really huge, so it is not so easy to retrieve information. I can spend hours searching for power point guide lines, without success. There is so much information which makes it hard to find the right one' (2, 8:20).

Respondent 3 on poor search engine implementation:

'It kind of doesn't guarantee that you find the most recent thing, it is not sorted in any way so it is sort of unreliable' (3, 9:43).

During the interviews respondents did express a strong animosity against the current system (and the organisation of information).

Respondent 9 on user assumptions:

'The problem is when it's been bad (red: the search engine and the intranet) it stays in people's minds that it's bad.' (9, 27:53)

However, it seems that this is partly based on assumptions and not in all cases on actual experience. The bias discovered among the respondents was taken into account when interpreting the outcome of interviews. A number of respondents did stress a strong use of person-to-person communication either by e-mails or telephone instead of using the information systems.

Respondent 3 on e-mail use:

'There is still a big culture around using e-mail. A lot will happen via e-mail and we have a huge amount of traffic still. There is also still that kind of

mentality where you feel that things are more formal and registered than just talking. But it does slow things down' (3, 7:36).

Respondent 2 on poor search engine performance and alternative communication:

'Usually it is quicker to ask and then receive the right information instead of searching for a solution' (2, 14:00)

To support results from the conducted interviews a walk through of selected parts of the WFPgo was undertaken. A walk through can be characterised as a systematic investigation (Kurg, 2000) in this case with the purpose of identifying consistency issues that can miss-lead users and subsequently result in users having difficulties in locating relevant information. A walk through can include both an evaluation of page layout and labels used in navigation menus. Here the so called trunk test defined by Steven Krug (2000) was used and a number of inconsistencies in the overall page layout and navigation system were found. This is in line with a statement collected from the interviews as expressed by respondent 3:

'And the content in general; I think we have gotten to a point where the content and links are just a bit too many. I mean the links on the left hand side which are supposed to be just the key ones have now become like a list of thirty' (3, 10:10)

The quote illustrates the general conception among users that the current information system is not supporting users in their quest for information, which might lead to the observed low perceived usefulness of the systems.

From the conducted users interviews and walk through of the WFP Intranet portal a number of issues with the present information architecture were identified. This includes a number of inconsistencies in the information architecture relating to the organisation of the portal and to a poor implementation of search engine facilities. As a result users expressed frustration with the current information architecture resulting in poor knowledge management and the use of external social networks for communication. Furthermore an extensive use of e-mail and telephone were observed when users were unable to retrieve information or documents from the intranet.

4.1 Search queries : log file analyses 1 (RQ1)

To qualify these finding and to further identify the search behaviour of WFP staff from both HQ and regional/local offices an analysis of the search query log from the Google Appliance Box was carried out. Previous research has demonstrated that a search query log can provide valuable information on users' information needs and how they transform these into a query formulation (Jansen, B. et al., 2000, Stenmark, 2006). A search log can also provide information on whether a user is making a navigational query instead of navigating to a page (Broder,

2002) that might be interpreted as a system having an illogical or confusing navigational structure.

In this study a search query log was analysed covering a three months period from 1st September to 30th November 2013. In total approximately 100.000 queries were recorded during this period and the average of terms pr. query is 1,8 words. The analysis proved that 13 out of the 25 top query terms were available as quick links on the front page of the WFPgo portal hence illustrating that among the top query terms a substantial number of queries can be categorized as navigational according to Broder (Broder, 2002) but also that the organization of the front page is not understood by users.

Search query log data were also used for evaluating the robustness of the proposed high-level taxonomy. The evaluation was conducted as a card-sorting exercise described later in this paper see 4.3.

4.2 Taxonomy (RQ2)

From previous studies (Gilchrist, 2001; Gilchrist, 2003; Chaudhry and Jiun, 2005; Broughton, 2006a) the implementation of a taxonomy is often suggested as a tool for securing a well-functioning information architecture. The dynamic nature of WFP as an emergency response organization requires design of an efficient information architecture and consequently also the structuring of a supporting taxonomy.

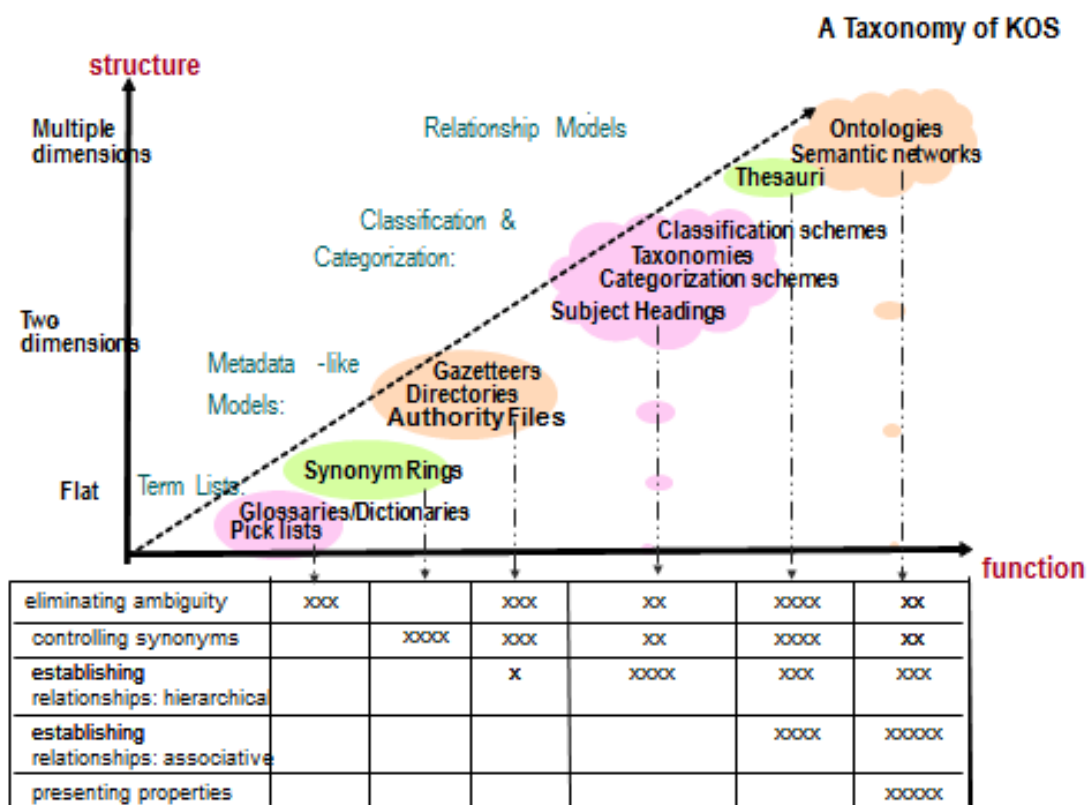


Figure 2. Structures and functions of various taxonomies and their benefits (Zeng, 2008, p. 161).

The illustration in figure 2 presents different and also very complicated Knowledge Organization Systems (KOS). Each type of KOS incurs different kinds of costs, each requires different levels of investment and some are easy to maintain while others are far more complex (Bedford, 2014). Hence, papers about implementation of KOS and a corporate taxonomy present problems as few researchers report about the subject.

The business literature shows that companies are concerned about being faced with an information architecture which does not help them in finding the documents they know exist. Eric Woods (2004) from Ovum, a research and consulting business on IT, observed that "Better information on what users are searching for, and the ability to tailor results and navigation paths, offers a relatively easy way to improve information retrieval within an organization" and he continues "A simple definition of a taxonomy is that it is a hierarchy of categories used to classify documents and other information. A corporate taxonomy is a way of representing the information available within an enterprise." However, no case studies implementing a corporate taxonomy were reported, probably because Woods in 2004 was the research director at the private company Ovum and thereby bound by company discretion.

In a research paper Sharma et al. wrote that "In creating corporate taxonomies, a practitioner makes use of metadata to describe documents and other resources thereby enabling a richer means of defining the context of the resource and to provide more information access points to support information query and retrieval operations." (Sharma et al., 2008, p. 37). The authors also provide a table for IT taxonomy tools but none of the links gave any information about implementations. It appears that Sharma et al. is restricted either by company confidentiality or that the intention of the paper is different from implementation a corporate taxonomy.

Pincher (2010b) shows another approach in an paper about operating a large-scale taxonomy project. He provides some simple rules saying, "Avoid the tendency to make the taxonomy more complex than it needs to be, you don't need to describe every possible category, just the ones needed today." Although Pincher does not offer any samples about implementation of the project, his down to earth approach is "keep the taxonomy simple" and the overall message is quite refreshing, however, the execution is a theoretical approach.

Bedford (2014) focuses on taxonomy assets or liabilities from an economic perspective and she summarizes the different KOS types. In her paper she stresses (ibid., p. 21) that "Arguments for enterprise taxonomies should be grounded in economics not just data or information management principles. Any decision to invest in taxonomies should be made with clear economic expectations". WFP has not asked for any economic models but in the presentation of an altered information architecture Bedford's approach about liabilities is significant when different alternatives are considered.

Based on the apparent and expressed problems in WFP and the few statements about implementation it was decided to present the taxonomy of KOS (figure 2) for the users. During the first workshop it was also suggested to utilise Hedden's broad framework and try out the least structured functions of the whole range of taxonomies i.e. a pick list (a controlled vocabulary) and synonym rings (Pellini and Jones, 2011b). Liability is taken into consideration as the least resource demanding functions are chosen.

At a minimum, a pick list is simply a restricted list of key words for some specialized purpose, usually for indexing, labelling, or categorizing. It is controlled because only terms from the list may be used for the subject area covered. If used by more than one person, it is also controlled in the sense that there is control over who may add terms to the list and when and how they may do it. The list may grow, but only under defined policies. The objective of a controlled vocabulary is to ensure consistency in the application of index terms, or labels to avoid ambiguity and the overlooking of information if the wrong search term is used (Hedden, 2010).

Synonym rings allow one to bring together synonyms – true synonyms, close synonyms, misspellings, predecessors, abbreviations, acronyms and so forth. The ring structures help to expand the meaning of the words that are being searched (Bedford, 2014).

The creation of a corporate taxonomy is often based on a combination of using automated technologies and human resources, typically information specialist with domain knowledge (White, 2007, p. 17) for deriving terms for use in the taxonomy. A intranet search engine can be considered an automated taxonomy creation tool since the possibility to apply filtering techniques to refine search results can be seen as a on-the-fly creation of a taxonomy like facilities (White, 2007; Russell-Rose and Tate, 2013). According to Gilchrist (2003) the resources for developing a corporate taxonomy incorporating a relationship model (see fig. 2) is high and in the case of WFP the human resources available for creating and maintaining a taxonomy is limited. Pincher (2010a) supports this by providing a firm statement saying that "The price of implementing a simple business taxonomy pales in comparison to the cost of failure of a project lacking one".

4.3 Development of High Level Taxonomy and card-sorting experiment (RQ2)

BCG has worked with WFP on Knowledge Management since 2012. They provided a High Level Taxonomy (HLT) which was presented for discussion at a workshop for WFP managers in October 2013. The HLT proposed by BCG was based on WFP's value chain (see figure 3) enumerating the various phases every emergency project traverses. Each of the four elements have different functions attached and the value chain is guiding the development of the proposed HLT.

The three functions utilized for the HLT are: Operational Functions,

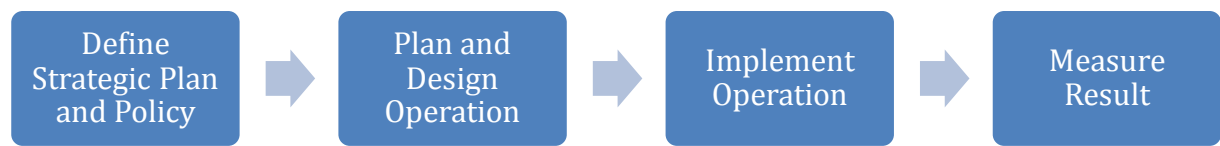


Figure 3. WFP value chain.

Support Functions and Themes and Cross Themes. The latter can change content and designations according to emergency projects in WFP. The adaptation of the different functions are done in the following manners: Operational Functions and the corresponding terms (see figure 4) have a specific a priori link to one of the phases in the value chain i.e. the expression “Strategic Planning” can only be attached to the first phase in the value chain, the terms “Procurement, Security and Logistic” can only be used in the second and third phase and the frase “Performance Management & Monitoring (incl Risk management)” is confined to the second, third and fourth phase. The expression “Operational Information Management & Reporting” is restricted to third and fourth phase while the term “Evaluation” can only be attached to phase four of the value chain.

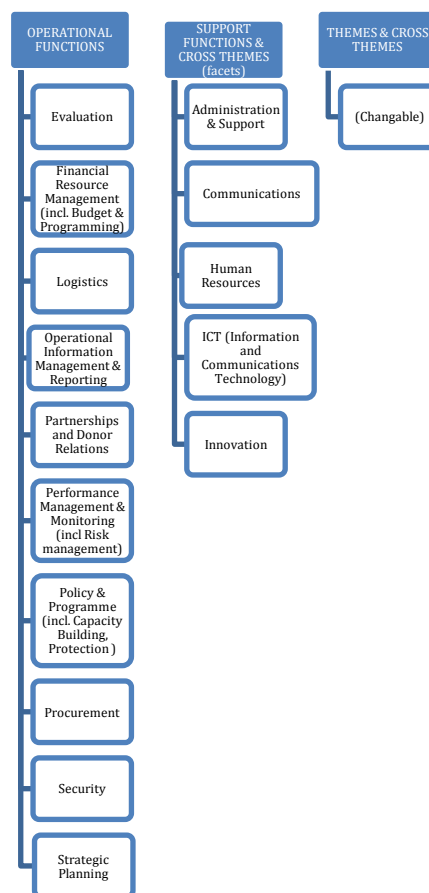


Figure 4 : First High Level Taxonomy (HLT).

In a paper presented in 2003 Bruno and Richmond provide a table about taxonomies representing information where the pros and cons are listed (table 1).

Types of Taxonomies	
Functional: This type of taxonomy organize itself along the different functions performed by an organization – both administrative and operational	
Pros <ul style="list-style-type: none"> • is most in tune with organizational goals and business processes • reduce silos of information • redices duplication • makes it easier to find the most recent official document • shows the flow of information • naming of headings is unaffected by department name • is recommended by ISO Technical Report 15489 	Cons <ul style="list-style-type: none"> • requires a new way of thinking about information • needs buy-in from everyone • requires one person to oversee the major shared “buckets” • requires a liaison within each department that contributes to the “bucket” • requires more training of employees

Table 1. Functional taxonomies (Bruno and Richmond, 2003, p. 48)

The taxonomy described as “Functional” the cons enumerate many of the difficulties experienced in WFP, and the pros register several benefits of a taxonomy based on functions. In WFP these pros signify a wish-list and they are dealt with in developing a HLT. However, even though we strongly agree with Bruno and Richmond the lack of an implementation model is a challenge.

Terms for Support Functions are evenly accepted under all four stages in the value chain. The terms are understood as facets i.e. the domain specific angle a function/theme is regarded from. This viewpoint is not fully supported by Broughton who argues that “Many introductory books on classification present faceted classification in a simplified but rather limited manner by using an example based on the attributes of entities” (Broughton, 2006b, p. 52). Hence, the unproblematic understanding is utilized because of the limited resources in WFP which make it necessary to cope with an unpretentious solution.

To get users’ view about the HLT proposed by BCG and its usefulness in structuring the search terms from WFPgo for navigational/informational Web queries 167 keywords and phrases were randomly selected from the log-file (query terms from 1st September – 29th November 2013) and a Card-sorting exercise was carried out. The card-sorting method is used to generate information about the associations and grouping of specific data items. Participants in a card sort are asked to organize individual, unsorted items into groups and may, depending on the technique, also provide labels for these groups. Card-sorting studies can provide insight into users’ mental models,

illuminating the ways they often tacitly group, sort, and label tasks and content in their own heads (Morville and Rosenfeld, 2007).

We used a closed card-sorting. The purpose of this card-sorting is to explore how new content may fit into an existing structure, or how new content relates to existing content. In a closed card-sort, participants are provided with a set of items to sort and pre-defined categories and are asked to place the items into the predefined categories. We provided the 167 cards and asked the participants to organize these into the pre-arranged HLT (under the value chain) that has been put on the table in front of them.

The outcome of the exercise showed that the staff didn't consider the value chain as superior to the HLT. Although different HLT terms (representing Operational Functions) had a priori link to various phases of the value chain the employees applied the card terms to any part of the HLT. The Support Functions, which were presented as facets to be utilized together with any of the HLT terms were all considered as HLT terms. This conduct indicated that the value chain had no meaning as a superior level to the HLT and that Support Functions as facets were not respected by the staff. The analysis showed that:

- the HLT could be employed to organize the keywords
- the value chain had no function in the taxonomy structure
- facets (Support Functions) should be included and treated as HLT terms

One issue was the geographical terms. The suggestion was to incorporate them as facets to be included in all groups in the HLT that is, any HLT term can be seen from any geographical view (facets) e.g. Administration & Support have the facets (is seen from) Thailand or Asia (figure 5). This function was acceptable by the staff.

(EXTENDED) HIGH LEVEL TAXONOMY (HLT)

Administration & Support
Communications
Evaluation
Financial Resource Management (incl. Budget & Programming)
Human Resources
ICT (Information and Communications Technology)
Innovation
Logistics
Operational Information Management & Reporting
Partnerships and Donor Relations
Performance Management & Monitoring (incl. Risk management)
Policy & Programme (incl. Capacity Building, Protection)
Procurement
Security
Strategic Planning
[Cross themes (examples)]
 Emergency Preparedness
 Gender

Facets

(to be added to any HLT group)

Geographical facets

- **Regions**
 - **Countries**
 - **Towns**
 - **Areas**

Example:

- **Asia**
 - **Thailand**
 - **Bangkok**

Figure 5. Extended categories in the High Level Taxonomy (HLT) and geographic facets.

4.4 Navigational and informational terms. : log file analyses 2 (RQ1 + RQ2)

For the organization of information, we often see taxonomies applied in Website information architecture (structural design), online information services, intranet content organization, and corporate content management systems. In such Website taxonomies, the emphasis is on guided user navigation rather than on search and retrieval of specific information. Navigation means finding one's way around (Lin et al., 2008) and in information science it is utilized as a metaphor where you determine your position and path through concepts and documents. The taxonomy for a Website is a lot like a table of contents, organized by topic. It can be reflected in the navigational menu and in the site map. As such, it might be called a navigational taxonomy (Hedden, 2010) (see figure 6).

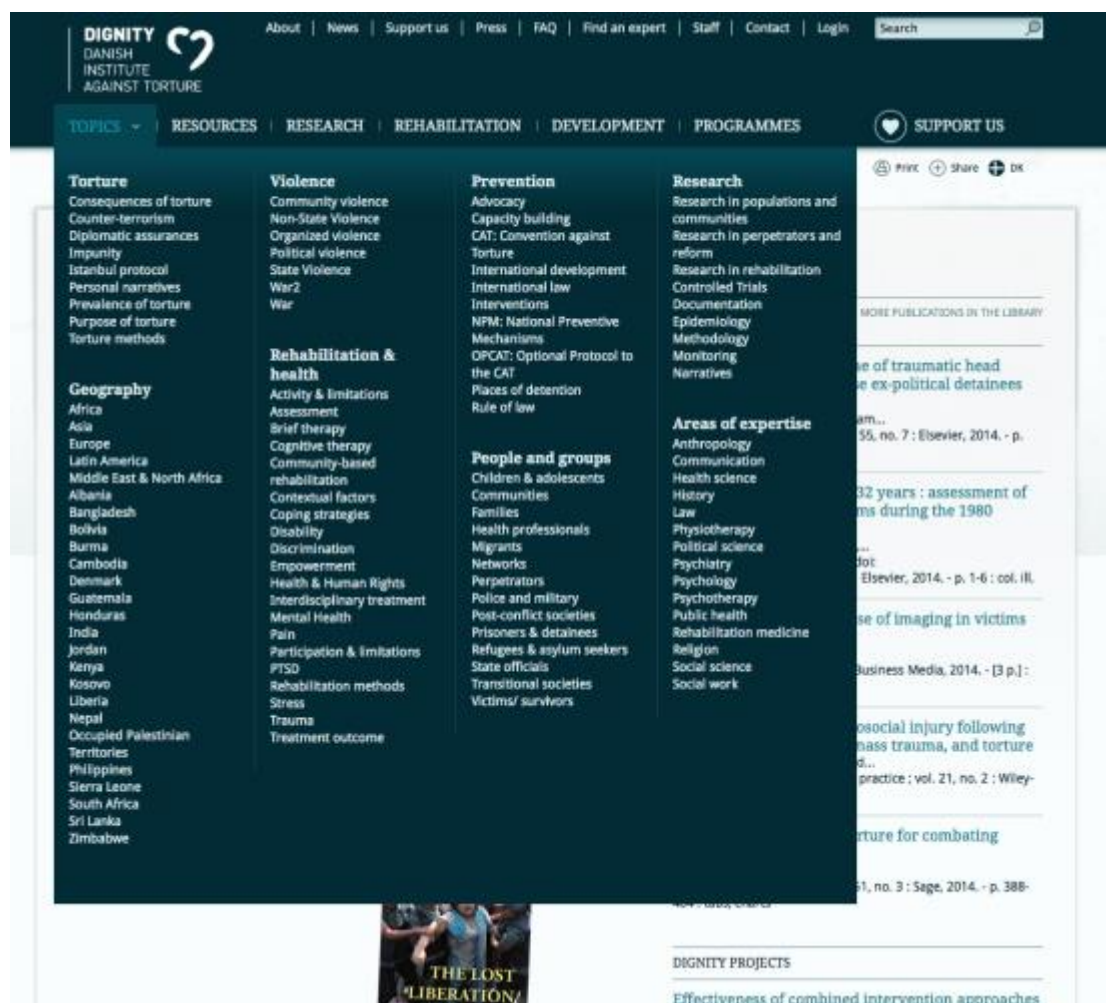


Figure 6. Topic organisation from the Danish Institute Against Torture (www.dignityinstitute.org)

We analysed the log-file from the Regional offices (queries made in September 2013) and divided the 1712 requests into navigational and informational terms. Broder (2002) classify Web queries according to their intent into 3 classes:

1. **Navigational.** The immediate intent is to reach a particular site.
2. **Informational.** The intent is to acquire some information assumed to be present on one or more Web pages.
3. **Transactional.** The intent is to perform some Web-mediated activity.

Broder's two first groups are of interest to WFP. The purpose of navigational queries is to reach a particular site that the user has in mind, either because they visited it in the past or because they assume that such a site exists. This type of search is sometimes referred as known item search. These types of queries are essentially navigational queries (ibid). This understanding is supported by Rosenfeld (2011) who, among his six content types based on query patterns has

one termed Navigational pages which concerns sub-sites where users just wanted to go to a site for a department or other program.

According to our log analysis and the above classification navigation is expressed as:

- 1) information needs that can be navigated directly to a site;
- 2) specific queries that indicate a known result e.g. site names like *COMET*; *compass*; and *EPIC* and
- 3) specific documents e.g. guides, check lists and manuals like *UN exchange rate June 2013*; *administrative calendar*; and *apppmr learning materials*.

The purpose of Broder's second class (ibid) Informational is to find information assumed to be available on the Web and not created in response to the user query. According to our analysis subject information is then:

- 1) terms which indicate that the user assumes this kind of information to be covered by the system.

We divided the subject information into generic and specific queries. The specific subject queries are formulated precisely and indicates a specific information need (example: accountability towards affected population, Are SC contract holders entitled to compassionate leave%3F, causes of acute malnutrition , Duty station category E).

The generic subject queries are formulated more general (example: eguide, external recruitment, improvement plan, lost way bill)

The generic terms used in the query can be employed in a pick-list. The Specific terms from the query can be utilized in synonym rings.

5. Maintaining the taxonomy (RQ2)

Pelini and Jones (2011b) express that taxonomies have to be flexible and adaptable to change within the organisation. Taxonomies require inspections of their usefulness at frequent periods. If a longer interval e.g. half a year the taxonomy will stay relevant with the procedures and functions within the company. In WFP a taxonomy manager can guarantee that taxonomies stay current. Upholding of the taxonomy has to be part of any taxonomy project, regardless of how it is employed.

Social networking is based on its success that the freedom the Internet provides, and is generating opportunities for new ways of creating and managing taxonomies. Tagging and social bookmarking seem to be the way taxonomy is moving forward (Rafferty and Hilderley, 2007). If organizations do not consider social tagging (un-controlled keywords) they are in denial of the new opportunities they provide (Springer et al., 2008).

The choice on how complicated the taxonomy has to be must consider the expenses for the organisation. To overcome information silos and better cross-company management, rivalling systems can co-occur with a taxonomy. The

choice is not between taxonomy and a crowd-driven classification via tags, but rather in their integration to maximize the opportunities they both provide (Pellini and Jones, 2011b).

5.1 Establishing synonym rings. (RQ2)

Synonym rings are a special type of controlled vocabulary merely utilized during searching. They cannot be employed in the indexing process as none of them are designated as the preferred term. According to Zeng (2005) "... synonym rings ensures that a concept that can be described by multiple synonymous or quasi-synonymous terms will be retrieved if any one of the terms is used in a search" see figure 2.

Retrieval systems can be improved with synonym rings where the subject matter is only indexing with a limited pick-list. As synonym rings are found in the interface to electronic system e.g. Google, a synonym ring may be generated automatically from clusters of co-occurring terms in full text or developed by subject specialists (Fayen, 2004).

In a hands-on exercise the users were asked to establish true synonym rings and associated synonym rings based on queries from WFPgo from September 2013. The participants agreed that synonym rings were very beneficial in searching as one didn't have to consider acronyms and their spelled out forms, related and alternative terms, associations etc. (figure 7a and 7b).

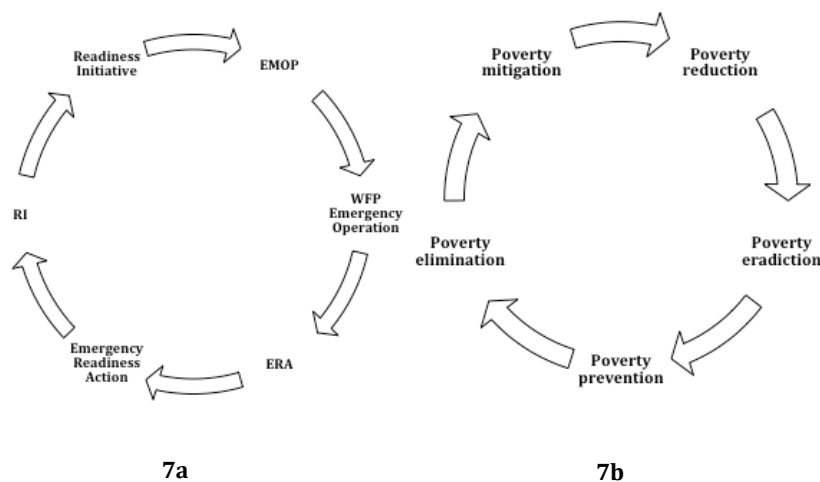


Figure 7a shows a synonym ring with (true synonyms) acronyms and spelled out forms and associated terms. **Figure 7b** shows a synonym ring with associated terms (Zeng, 2008).

During 2013 WFP managed immense emergency works in Syria and the Philippines. We analysed the queries from the log-files both from HQ in Rome and in the two country offices in Syria and the Philippines to perceive if synonym rings could be established on a permanent basis for emergency work and other kind of relief efforts. However, nothing indicated from the log-files that synonym rings with a specific emergency terminology were required. The reason is perhaps that WFP's aim is to deliver emergency food aid to victims of natural

disasters, war and civil conflicts. Therefore the terminology in the organization is already geared towards this goal. It seems that synonym rings for catastrophes in this context are redundant, hence, if the demand rises they can certainly be created.

6. Organization of future taxonomy work

It is evident that different categories of staff in WFP have different knowledge and information needs; thus each staff group should be able to access knowledge according to their specific needs and through the most appropriate means for each category (Oliver, 2008). However, without coordination nobody will be able to identify and address the potential knowledge gaps existing between the needs of staff and the knowledge available within the organization.

We suggested to and discussed with the WFP staff to have a taxonomy manager responsible for the preparations and maintenance of the taxonomy (pick-list) and for running and controlling the synonym rings. Additionally we recommended the setting up of taxonomy champions in each division and proposed the founding of a taxonomy network.

Implementation of the taxonomy and delivery of outputs need to be managed by a team with the necessary skills and expertise as well as access to key actors. The taxonomy champions shall analyse the needs and perspectives of all users. The taxonomy manager shall be in charge of maintaining the corporate taxonomy within the organization. There is no literature recommending the total of pick-list terms, however, we suggest 6-800 terms in addition to the 18 top-level categories which is a bit more than Pincher (2010a) recommends. Our advice is based on projects where the experience is that a restricted number of terms are controllable by a taxonomy manager. Depending on the complexity of the taxonomy project Pincher (ibid.) suggests that the champions' number should be between 6 and 12 members, with various expertise areas. Another key criterion for selection is sufficient seniority to access senior management level for key decisions and approvals (ibid).

Taxonomies can be employed to require all staff to index documents uploaded to the intranet. However, all actors must be advised as to the reason of the taxonomy, development and what is anticipated from the employees. In WFP the staff is already uploading documents to the system indicating that the employees are handling metadata although the term is unfamiliar to them. Comments from the last workshop indicated, however, that the altered indexing using a pick list and synonym rings could easily substitute the somewhat unsystematic categorization which is currently exercised.

7. Conclusion

An efficient information architecture is an imperative for users' accepting and using a corporate intranet portal. As shown in this case study users tends to find alternative solutions for knowledge management and document access if they encounter difficulties in understanding the organization of a corporate intranet. Alternative solutions for information and document access may lead to an extensive use of person-to-person communication by e-mail or telephone and the employment of information systems outside the control of the organization. This includes the use of social networks and file sharing services. A tool for achieving an efficient information architecture is to develop a corporate taxonomy. In this study the analysis of the existing information architecture was exposed through staff interviews and walkthroughs of the intranet portal. A number of the problems identified clearly showed that development and implementation of a taxonomy could minimize the drawbacks within the portal navigation and the searching for information and documents using the available search engine.

It is suggested to support the dynamic nature of the WFP as an emergency response organization with a high level taxonomy where the number of terms are limited and supplemented by a rather flat structure based on a pick-list and the use of synonym rings for synonyms and associated terms.

With only limited resources available for taxonomy development within the organization, we believe that the only possible solution is to involve end users directly in the development of the pick list and synonym rings. This also guided the decision not to suggest the development of a taxonomy based on a relationship model but instead choose the less resource demanding solution. To support the findability of documents it is then important that the organization is forward thinking in implementing an efficient intranet search solutions with taxonomy like filtering techniques.

The proposed taxonomy has not been evaluated yet, but this must be seen as an import next step in gaining a more profound insight into the proposed solution. An evaluation phase has been suggested to WFP by the authors.

The search log analysis and card-sorting exercises gave valuable insights about how WFP staff interpreted the HLT and led to a number of changes of the taxonomy structure. We believe that this blended approach based on the value-chain cater for a solid foundation of the proposed taxonomy and to the knowledge of the authors are an approach that has not been thoroughly described before in the literature.

In order to continue the development of the taxonomy it is suggested that WFP formalizes the work through a taxonomy manager who shall rely on taxonomy champions appointed from each departments securing the commitment in the further taxonomy work by including the taxonomy end users. The recommendations have been implemented by WFP in 2015.

Organizations that use knowledge to facilitate their processes and to improve the staff's understanding and development of a taxonomy tool tend to grow. The more closely the knowledge organization methods match the use of knowledge by the organization, the more efficiently the application and access of the knowledge. However, a future, thorough evaluation is compulsory and will provide a deeper understanding of the present and future knowledge requirements in WFP.

References

- ANSI/NISO Z39.19 (2005), "Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies"., available at: <http://www.niso.org/kst/reports/standards/> (accessed 27 August 2015).
- Bedford, D. A. D. (2014), "Understanding and managing taxonomies as economic goods and services", *Bulletin of the Association for Information Science and Technology*, Vol. 40 No. 4, pp. 15-22, available at: <http://onlinelibrary.wiley.com/doi/10.1002/bult.2014.1720400410/epdf> (accessed 22 November 2015).'
- Broder, A. (2002), "A taxonomy of web search", *SIGIR Forum, Fall 2002*, Vol. 36 No. 2, available at: <http://sigir.org/files/forum/F2002/broder.pdf> (accessed 27 August 2015).
- Broughton, V. (2006a), *Essential thesaurus construction*, Facet, London.
- Broughton, V. (2006b), "The need for a faceted classification as the basis of all methods of information retrieval", *Aslib Proceedings: New Information Perspectives*, Vol. 58 No. 1/2, pp. 49-72.
- Bruno, D. and Richmond, H. (2003), "The truth about taxonomies", *Information Management Journal*, Vol. 37, No. 2, pp. 44-53
- Chaudhry, A.S. and Jiun, T.P. (2005), "Enhancing access to digital information resources on heritage", *Journal of Documentation*, Vol. 61 No. 6, pp. 751-776.
- Conway, S. and Sligar, C. (2002) *Unlocking knowledge assets*, Microsoft Press, Redmond, Wash.
- Davis, F. D. (1989), "Pervieved Usefulness, Percieved Ease of Use, and User Acceptance", *MIS Quarterly*, 13(3), 319-340.
- Fayen, E. (2004), "A new standard for controlled vocabularies", *The Indexer*, Vol. 24 No. 2, pp. 62-65.
- Fenwick, T., Seville, E. and Brunsdon, D. (2009), *Reducing the Impact of Organisational Silos on Resilience : A Report on the impact of silos on resilience and how the impacts might be reduced*, Resilient Organisations, New Zealand.
- Findwise (2014) "Enterprise search and findability survey", available at <http://www.findwise.com/findabilitysurvey2014> (accessed 25 November 2015)
- Garshol, L. M. (2004), "Metadata? Thesauri? Taxonomies? Topic Maps! Making Sense of it all", *Journal of Information Science*, Vol. 30 No. 4, pp. 378-391.

Gilchrist, A. (2001), "Corporate taxonomies: report on a survey of current practice", *Online Information Review*, Vol. 25 No. 2, pp. 94 -103,

Gilchrist, A (2003), "Thesauri, taxonomies and ontologies – an etymological note", *Journal of Documentation*, Vol. 59 No. 1, pp. 7-18.

Hedden, H. (2010), *The Accidental Taxonomist*, Information Today Inc., Medford, NJ.

Jansen, B. J., Spink, A. and Saracevic, T. (2000), "Real life, real users, and real needs: A study and analysis of user queries on the web", *Information Processing and Management*, Vol. 36 No. 2, pp. 207-227.

Krug, S (2000), *Don't make me think! A common sense approach to web usability*, New Riders Publishing, Berkeley, CA.

Kvale, S. and Brinkmann, S. (2009), *Interviews: An Introduction to Qualitative Research Interviewing*, Sage Publications, Thousand Oakes, CA.

Lambe, P, (2007), *Organising Knowledge: Taxonomies, Knowledge and Organisational Effectiveness*, Chandos Publishing, Oxford.

Lambe, P. (2011), "The Unacknowledged Parentage of Knowledge Management", *Journal of Knowledge Management*, Vol. 15 No. 2, pp. 175-197.

Lin, J., DiCuccio, M., Grigoryan, V. and Wilbur, W.J. (2008), "Navigating information spaces: A case study of related article search in PubMed", *Information Processing and Management*, Vol. 44 No. 5, pp. 1771–1783.

Morville, P. and Rosenfeld, L. (2007), *Information Architecture for the World Wide Web*, 3rd Ed. O'Reilly Media, Inc., Sebastopol, CA.

Oliver, G. (2008), "Information culture: exploration of differing values and attitudes to information in organisations", *Journal of Documentation*, Vol.64 No. 3, pp. 363-385.

Pellini, A. & Jones, H. (2011a), *A Study of ADB's Knowledge Taxonomy. Final Report*, Knowledge Management Center. Regional and Sustainable Development Department, Manila, ,available at: <http://www.adb.org/sites/default/files/pub/2011/knowledge-taxonomy.pdf> (accessed 27 August 2015).

Pellini, A. & Jones, H. (2011b), *Knowledge taxonomies. A literature review*, Overseas Development Institute (ODI), London,, available at: <http://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/7123.pdf> (accessed 27 August 2015).

Pincher, M. (2010a), "A Guide to Developing Taxonomies for Effective Data Management", *Computer Weekly*, 8 March available at: <http://www.computerweekly.com/Articles/2010/04/06/240539/A-guide-to-developing-taxonomies-for-effective-data.htm> (accessed 27 August 2015).

Pincher, M. (2010b), "Case study: compliance with complex construction and multi-disciplinary teams", *Computer Weekly*, 8 March, available at: <http://www.computerweekly.com/feature/Case-study-compliance-with-complex-construction-and-multi-disciplinary-teams> (accessed 22 November 2015).

Rafferty, P. and Hilderley, R. (2007), "Flickr and Democratic Indexing: dialogic approaches to indexing", *Aslib Proceedings: New Information Perspectives*, Vol. 59 No. 4/5, pp. 397-410.

Rosenfeld, L. (2011), *Search Analytic for your Site. Conversation with Your Customers*, Rosenfeld Media, New York, NY.

Russell-Rose, T & Tate, T (2013) *Designing the search experience : the information architecture of discovery*. Morgan Kaufmann, Amsterdam

Sharma, R.S., Foo, S. & Morales-Arroyo, M. (2008), "Developing corporate taxonomies for knowledge auditability: A framework for good practices", *Knowledge Organisation*, Vol. 35 No. 1, pp. 30-46.

Springer, M., Dulabahn, B., Michel, P., Natanson, B., Reser, D., Woodward, D. and Zinkham, H. (2008), *For the Common Good: The Library of Congress Flickr Pilot Project*, Library of Congress, Washington, DC., available at: http://www.loc.gov/rr/print/flickr_report_final.pdf (accessed 27 August 2015).

Stenmark, D. (2006), "Intranet users' information-seeking behaviour: an analysis of longitudinal search log data", *Proceeding of the American Society for information science and technology*, Vol. 43 No. 1, pp 1-19.

Stocker, A., Zoier, M., Softic, S., Paschke, S., Bischofter, H. and Kern, R. (2014), "Is enterprise search useful at all? : lessons learned from studying user behaviour", in *i-know '14. Proceeding of the 14th International Conference on Knowledge Technologies And Data-driven Business*, ACM.

Tariq, A., Georgiou, A. and Westbrook, J. (2014), "Coping with information silos: An examination of the medication management process in residential aged care facilities (RACFs)", in Grain, H. et al. (Eds) *Investing in E-Health: People, Knowledge and Technology for a Healthy Future*, IOS press, Amsterdam, pp. 156-162, available at: <http://ebooks.iospress.nl/publication/37247> (accessed 25 November 2015).

Taxonomies & Controlled Vocabularies SIG (2013). "A Special Interest Group of the American Society for Indexing" available at: <http://www.taxonomies-sig.org/about.htm#taxonomy> (accessed 27 August 2015).

Wang, S. and Noe, R.A. (2010), "Knowledge sharing: a review and directions for future research", *Human Resource Management Review*, Vol. 20, pp. 115-131.

Whittaker, M. and Breininger, K. (2008), "Taxonomy Development for Knowledge Management", in *World Library and Information Congress: 74th IFLA General Conference and Council. Québec, 10-14 August 2008*, pp. 1-10, available at: http://archive.ifla.org/IV/ifla74/papers/138-Whittaker_Breininger-en.pdf (accessed 27 August 2015).

White, M. (2007), *Making search work: implementing web, intranet and enterprise search*, Facet, London.

White, M. (2010), "Intranet management: divine comedy or strategic imperative?", *Ariadne*, No. 62, available at: <http://www.ariadne.ac.uk/issue62/white> (accessed 4 December 2015)

Woods, E. (2004), "The corporate taxonomy: creating a new order", *KMWorld*, Vol. 13 No. 7, available at: <http://www.kmworld.com/Articles/Editorial/Features/The-corporate-taxonomy-creating-a-new-order-9566.aspx> (accessed 23 November 2015).

World Food Programme (2014), *All about the World Food Programme*, available at: <http://documents.wfp.org/stellent/groups/public/documents/newsroom/wfp261411.jpg> (accessed 27 August 2015).

Zeng, M.L. (2005), "Construction of Controlled Vocabularies, A Primer (based on Z39.19)", Available at: <http://www.slis.kent.edu/~mzeng/Z3919/index.htm> <http://marciazeng.slis.kent.edu/Z3919/3structure.htm> (accessed 27 August 2015).

Zeng, M.L. (2008), "Knowledge Organization Systems (KOS)", *Knowledge Organization*, Vol. 35 No. 2/3, pp. 160-182.

