Agarwal, N.K. & Poo, D.C.C. (2006). Meeting knowledge management challenges through effective search. *International Journal of Business Information Systems, 1*(3), 292-309. https://doi.org/10.1504/IJBIS.2006.008601

Meeting knowledge management challenges through effective search

Naresh Kumar Agarwal* and Danny C.C. Poo

School of Computing
National University of Singapore
3 Science Drive 2, 117543 Singapore
E-mail: naresh@comp.nus.edu.sg
E-mail: dpoo@comp.nus.edu.sg
*Corresponding author

Abstract: With the emergence of information- and knowledge-based economies, knowledge management has gained importance in organisations. It remains a challenge, with a typical mid-sized organisation quickly accumulating vast amounts of information. Trying to make sense of data and to search effectively forms a core part of this challenge. This paper aims to demonstrate that information providers and information finders must work hand-in-hand to achieve a 'Fit', leading to successful information finding by one of four categories of roles that finders assume. It is proposed that this 'Fit' can be achieved by combining a few or more of the following means: 1) taxonomy-based presentation/classification; 2) multiple views of the taxonomy; 3) personalisation; 4) usage of cues; and 5) usage of semantics. A search from the internet would also benefit from 6) localisation and 7) specialty search. Improved search and navigation arising from the Fit should enhance the efficacy of knowledge management in an organisation.

Keywords: knowledge management; information retrieval; search; taxonomy; personalisation; cues; semantics; localisation; specialty search.

Reference to this paper should be made as follows: Agarwal, N.K. and Poo, D.C.C. (2006) 'Meeting knowledge management challenges through effective search', *Int. J. Business Information Systems*, Vol. 1, No. 3, pp.292–309.

Biographical notes: Naresh Agarwal is Research Assistant and Graduate Student in Information Systems at the School of Computing, National University of Singapore. He has over five years of diverse industry experience, principally in the design and development of solutions that helped fill technology gaps and enhance business functions. Agarwal holds a Bachelor of Applied Science in Computer Engineering (Honours) from the Nanyang Technological University, Singapore. His current area of research is information management, classification and retrieval.

Dr. Danny Poo is tenured Associate Professor and Assistant Dean at the School of Computing, National University of Singapore. He holds a BSc (Honours), MSc and PhD in Computation from the University of Manchester Institute of Science and Technology, UK. He has extensive experience in software engineering, information management and knowledge management, both in the technical and managerial levels. He has carried out wide-ranging research in the areas of effective search strategies, relevance ranking and feedbacks, electronic referencing, meta-data, taxonomy generation, information sharing,

knowledge management portals, object-oriented software engineering and software systems maintenance. He has served in the Steering Committee of the Asia Pacific Software Engineering Conference (APSEC) since 1994 and was recently appointed Vice-Chairman of APSEC. He is an author of three books on Java/J2EE.

1 Introduction

1.1 Emergence of knowledge- and information-based economies

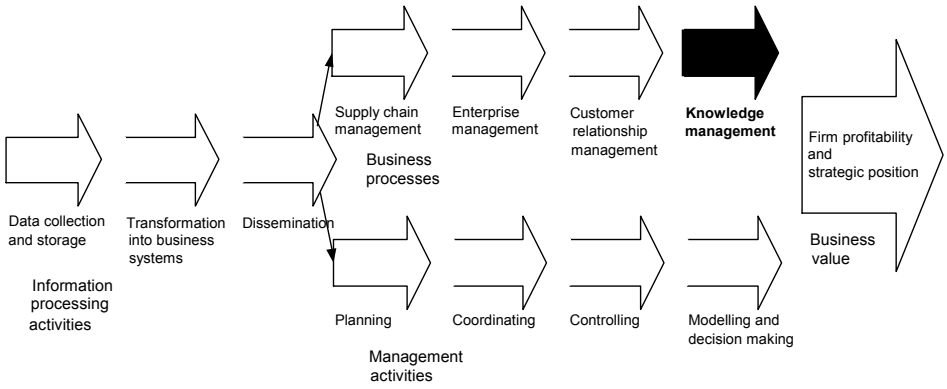
Businesses today are undergoing major changes with information systems helping companies extend their reach to faraway locations, offer new products and services, reshape jobs and work flows, and perhaps profoundly change the way they conduct business. Information technology constitutes more than 70% of the invested capital in service industries such as finance, insurance and real estate.

A global economy has emerged and become strong. Industrial economies and societies are being transformed into knowledge- and information-based service economies, where knowledge and information are key ingredients in creating wealth. Today, most people no longer work on farms or in factories but instead are found in sales, education, healthcare, banks, insurance firms and law firms; they provide business services like copying, computer programming, or making deliveries. These jobs primarily involve working with, distributing, or creating new knowledge and information.

The business enterprise is transforming[1] itself into a flattened (less hierarchical), decentralised, flexible arrangement of generalists who rely on nearly instant information to deliver mass-customised products and services uniquely suited to specific markets or customers. The new manager relies on informal commitments and networks to establish goals (rather than formal planning), a flexible arrangement of teams and individuals working in task forces, and a customer orientation to achieve coordination among employees. He/She appeals to the knowledge, learning and decision making of individual employees to ensure proper operation in the firm (Laudon and Laudon, 2004).

All this is leading to the emergence of the digital firm where nearly all of the organisation's significant business relationships with customers, suppliers and employees are digitally enabled and mediated. Along with supply chain management systems, customer relationship management systems and enterprise systems, knowledge management systems (see Figure 1) is a core area where corporations are digitally integrating their information flows (Laudon and Laudon, 2004).

Figure 1 The business information value chain. From a business perspective, information systems are part of a series of value-adding activities for acquiring, transforming and distributing information that can be used to enhance decision making and performance, and ultimately increase firm profitability. Knowledge Management finds an important place.



Source: Adopted from Laudon and Laudon (2004)

1.2 The knowledge management problem – retrieving relevant information

A collection of data is not information, a collection of information is not knowledge, a collection of knowledge is not wisdom and a collection of wisdom is not truth (observation by Fleming, 1996). Knowledge Management plays an important role in an organisation by facilitating the capture, storage, transformation and dissemination of information. Organisations have been experimenting with knowledge management to try and promote efficiency, improve profits, be innovative and have a competitive advantage and, sometimes, simply to survive (Wigg, 1997; Prusak, 1997; Hendriks and Virens, 1999; Loucopoulos and Kavakli, 1999; Davenport and Prusak, 2000; Gao *et al*., 2002).

Managers have encountered problems when attempting to transform their firms through Knowledge Management programs (Gold *et al*., 2001). Information systems that truly enhance the productivity of knowledge workers may be difficult to build because the manner in which information technology can enhance higher-level tasks, such as those performed by managers and professionals, is not always clearly understood. Some aspects of organisational knowledge cannot be captured easily or codified, or the information that organisations finally manage to capture may become outdated as environments change. It is very difficult to integrate knowledge management programs into business strategy. Processes and interactions between information technology and social elements in organisations must be carefully managed (Laudon and Laudon, 2004; Davenport *et al*., 2002; Grover and Davenport, 2001).

Knowledge Management, thus, remains a challenge, with a typical mid-sized organisation today accumulating more information in a week than the whole of human kind did between the years one and 1500 A.D. Masses of documents, e-mails, databases, images, audio and video recordings form vast repositories of information assets to be tapped by employees, partners, customers and other stakeholders (Papadopoullos, 2004). Trying to make sense of all this data and being able to effortlessly retrieve relevant information form a core part of Knowledge Management.

1.3 Knowledge base of an organisation

The knowledge base of an organisation may include:

- structured internal knowledge (explicit knowledge), such as product manuals or research reports
- external knowledge of competitors, products, and markets, including competitive intelligence
- informal internal knowledge, often called tacit knowledge, which resides in the minds of individual employees but has not been documented in structured form (Davenport *et al*., 1998).

Knowledge management systems try to identify, capture, codify and distribute this knowledge. All this forms part of the *knowledge that resides within and often belongs to the organisation*. Apart from these, a lot of employees in an organisation rely on the internet to search for information related to their work. This information forms potential *knowledge that resides outside the organisation* and is accessible through search engines.

1.4 Inadequacies of searching techniques in meeting search quality

There have been continual efforts to improve user-experience and information-retrieval through quality search and search-optimisation techniques (Shapiro and Lehoczky, 2003; Sullivan, 2005; Notess, 2004; Clough, 2005). However, a particular search engine may not be effective in meeting all the needs of a particular searcher, affecting search quality as a result. Lawrence and Giles (1998) conducted a study of six World Wide Web search engines and concluded that the coverage of any one engine is significantly limited: No single engine indexes more than one-third of the 'indexable web'.

With organisations accumulating huge amounts of data within short spans of time, retrieving information that resides within and belongs to the organisation is no mean task either.

2 Research goals

The authors of this paper have formed a number of hypotheses pertaining to information retrieval in an organisation. In this paper, the authors present a discussion on these hypotheses. Experiments will be conducted and empirical findings will be shared in subsequent papers.

Hypothesis 1 *Employees assume one of four roles when retrieving information from their organisation's knowledge base or the internet – that of the i) learner ii) data gatherer iii) location seeker or iv) focused searcher.*

Hypothesis 2 *Information providers (people in the organisation responsible for identifying, capturing and codifying knowledge or content providers via their search engines) and information finders must work hand-in-hand in order to achieve a 'Fit' leading to successful information finding by one of the four categories of searchers.*

Hypothesis 3 *The Information Provider-Finder Fit can be achieved by combining a few or more of the following means: i) taxonomy-based presentation/classification; ii) multiple views of the taxonomy; iii) personalisation; iv) usage of cues; and v) usage of semantics. The usefulness of information accessed from the internet would also improve through vi) geographical localisation and vii) specialty search.*

3 Four roles of information searchers in an organisation

Access to vast stores of information can no longer be ensured by simply launching the browser, going to one's search page (either belonging to the organisation or a general purpose search engine) and typing a word or two in the hope of locating one of the oft-required needles from the haystack. Search needs far exceed the simple requirement to locate a document with a given word in it (Papadopoullos, 2004). Search and classification results must satisfy four basic categories of users; those:

1 who need information about a topic they are not familiar with in preparation for starting a new project

2 who need information about a topic they *are* knowledgeable about and are therefore in data-gathering mode

3 who have a good idea what they are looking for, know that a given document or piece of data exists, and simply need to locate it

4 who need a very specific answer to a specific question (Papadopoullos, 2004).

We term these four categories of users as:

1 learner

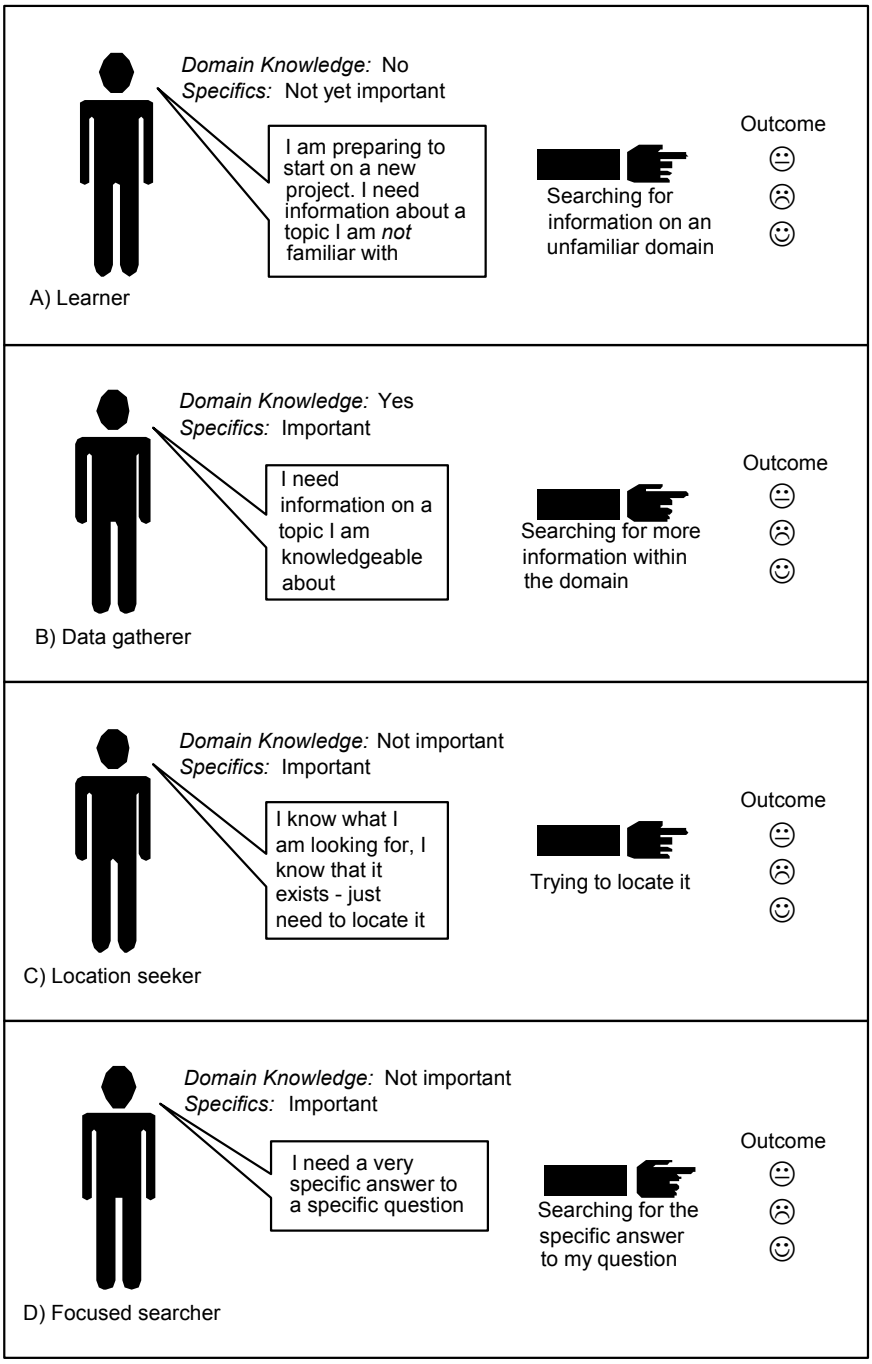
2 data gatherer

3 location seeker

4 focused searcher.

Depending on the context of the data one is searching for and the domain knowledge the person has in the field pertaining to the search, the same person may assume one of the four aforementioned roles (see Figure 2).

Figure 2 Four roles assumed by information seekers in an organisation



Note: Information finder = successful searcher = ☺

Information finder

In Figure 2, we have used emoticons to denote the satisfaction level of a searcher in an organisation. Depending on the query results, the searcher may be moderately satisfied with the search results 😐, dissatisfied with the search results ☹, or totally satisfied with the search results ☺. Based on a successful searcher ☺, we propose a definition for the Information Finder as follows: "*Information finder* is a role a searcher assumes when he/she is totally satisfied with the search results". The search results an information finder gets are relevant and closely match what the finder expected to find while initiating the search.

Information search service providers and content providers the world over are working to create more information finders at all times out of searchers of information, irrespective of whether the searcher is performing the role of a learner, data gatherer, location seeker or focused searcher, and regardless of whether the information being searched resides within the organisation or on the public internet.

4 The Information Provider-Finder Fit (IPFF) framework

While general purpose search engines are widely prevalent, their focus on providing a 'one-size-fits-all' model for a search leads to inadequate search results (van Ossenbruggen and Hardman, 2002), not meeting the differing needs of the different roles each searcher assumes at different times while searching for different pieces of information. Data within organisations is also not specifically geared to meet the different demands of these roles.

With information searchers not playing a very active role in helping search service providers better provide information, the needs of information finders are not adequately met. This gap (whether it be big or small, depending on context) between the search services provided and the needs of the finder cannot be bridged unless both the information provider and the information finder work hand-in-hand.

Inspired by the Task-Technology Fit Framework (Daft and Lengel, 1986; Zigurs and Buckland, 1998), an IPFF framework for information search and retrieval is proposed (as shown in Figure 3).

According to the framework, in order to achieve effective information finding, there must be a 'Fit' between the exact needs of the Information Finder and the services provided by the Information Provider (coupled with an exact understanding of the searcher's needs).

This would, of course, be the ideal case and is easier said than done. The establishment of this 'Fit' requires a departure from the current 'one-size-fits-all' model and the customisation of search results based on the specific needs (and an understanding of the knowledge level in the area of search) of the different searcher roles.

People may have a hard time stating why they are doing the actions that they are doing in an information-seeking task. The knowledge repository builder or the search service provider must provide mechanisms for the searcher to determine and specify the role he/she fits into while carrying out the search. On getting and knowing the required

context, the search engine must have mechanisms to work interactively with the searcher, until the searcher retrieves the search results relevant to his needs. This must happen in a manner and time frame which leaves the searcher satisfied and changes him/her from a searcher to a finder.

Figure 4 summarised the hypothetical interaction involved between the searcher and the provider in order to establish a fit for the different searcher roles.

Figure 3 Information Provider-Finder Fit framework

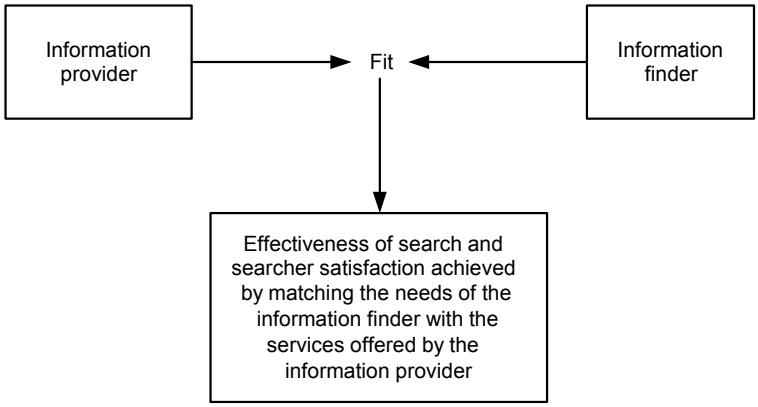
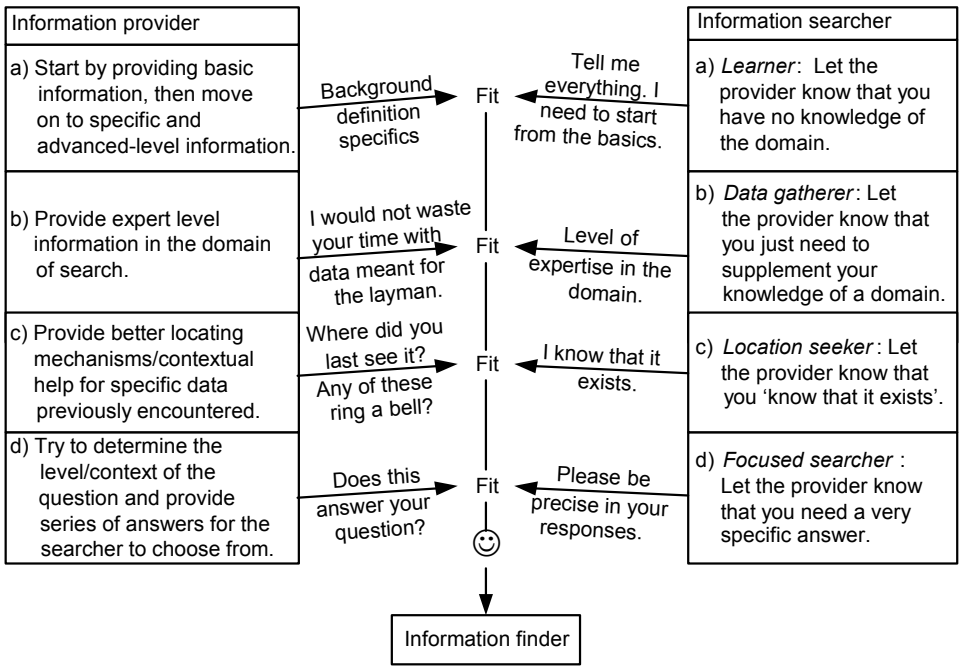


Figure 4 IPFF in the context of different roles searchers assume



A learner must let the provider know that he/she has no prior knowledge of the domain under search. The repository builder/search service provider must accordingly provide basic information, and then move on to more specific and advanced information. When a data gatherer is looking for information, a provider ought not to waste his/her time with basic information meant for the layman. It must provide mechanisms to determine the searcher's level of expertise in the domain and provide search results pertaining to the required level. A location seeker just needs to locate information previously encountered. The provider must provide locating mechanisms and contextual help. Similarly, the search service provider must rise up to the needs of the focused searcher who needs a specific answer to a specific question.

5 Mechanisms to achieve IPFF

The Knowledge Management needs of an organisation cannot be adequately met simply by trying to locate a piece of information in times of need. As Papadopoullos (2004) states, there is a need to move beyond "finding the needle in a haystack" to "connecting the dots" among various pieces of information.

Instead of adopting a 'one size fits all' information retrieval model, we propose a number of mechanisms that can be applied to knowledge repositories to help achieve a 'Fit' between Information Providers and Information Finders within an organisation: a) taxonomy-based presentation and classification; b) providing multiple views of the taxonomy; c) personalisation; d) usage of cues; and e) usage of semantics. When there is a need to access data from the internet (which is often enough), organisations would benefit if search engines provided f) localisation and g) specialty search, in addition to the five attributes mentioned above.

Using a combination of these mechanisms would provide a better 'Fit' for the different categories of roles searchers assume. What must be included for which category of users to achieve the best Fit for them will form the focus of subsequent research. Nevertheless, the authors provide their assumptions of usefulness in Section 6.

5.1 Mechanisms that aid the effective retrieval of knowledge residing in an organisation, as well as information accessible from search engines

5.1.1 Taxonomy-based presentation and classification

Taxonomies enable navigable search results that help users browse rather than search. They provide a subject-based classification that arranges the terms in a controlled vocabulary into a hierarchy. This allows related terms to be categorised into meaningful frameworks that add some logical structure that employees can rapidly navigate to find high concentrations of topic-specific, related information (Papadopoullos, 2004; Garshol, 2004). In an organisation, taxonomies could be organised based on departments, tasks at hand, projects, knowledge domains, areas of responsibility, *etc*.

5.1.2 Multiple views of the taxonomy

Taxonomies are flexible structures, and can be developed to cover many different topics to any desired level of granularity. Even more powerful are dynamic classifications that allow search results to be organised in real-time into classification views that are selected by the user in order to view information from various perspectives (Papadopoullos, 2004). Similarly, faceted classifications work by identifying a number of facets[2] into which the terms are divided, *e.g.*, classifying by colour, geography, subject, *etc*. Wilson (2003) uses a demo to classify wines into different facets – variety, region and price. A knowledge repository in an organisation could provide a view of information based on project, timeline, task or domain. Multiple tags could be associated with a single object that would enable it to be classified under different taxonomy views that are created on the fly, depending on the navigation pattern of the employee during a particular browsing/searching session.

5.1.3 Personalisation

Knowledge repositories need to be customised to an employee's individual needs. Personalisation would mean that the employee's profile, prior domain knowledge and expertise in the area of search (the particular role he/she is currently assuming) would be taken into account by the repository before giving out a set of results for a given query. Web search engines could also be similarly customised for each searcher role. Search results would then be best suited for each individual's need at the time of search.

However, this would imply that two employees, *A* and *B*, may or may not get the same results (or in the same order/ranking) when searching for the same keyword. This may cause a feeling of disorientation, especially when people collaborate on group projects. Thus, a general search option should be provided. Personalisation may also demand more time from the user, so it must be kept as simple as possible.

5.1.4 Usage of cues

The notion of context has been introduced to enhance search tools and refers to a diverse range of ideas from specialty-search engines to personalisation. It has been found that incorporating such contextual cues from static content sources, dynamic content sources, static collaborative sources and dynamic collaborative sources can help to increase the relevance of information (Goh *et al*., 2004) and enhance search quality by increasing the set of relevant results and decreasing non-relevant results.

In a knowledge repository in an organisation, static content sources such as names, designations and departments can be filled in by various departments. Dynamic content sources can be captured using a system that logs the users' actions. Users of the system can create a record of other team members so as to utilise the contextual cues that can be obtained from static collaborative sources. With such information, dynamic collaborative sources can also be obtained by matching the actions of the users with those of other team members with similar interests or work responsibilities. The contribution of each type of contextual cues will help to improve the search quality of the system.

5.1.5 Usage of semantics

While searching for information, an employee may not use the right keyword or phrase that is understood well by the search service provider. This would especially be true for the learner and location seeker. This leads to inadequate search results. The efficacy of search can be enhanced by the usage of semantics, language-based constructs and thesauri-based services for better keyword matching (Martin and Eklund, 1999; Cao *et al*., 2003).

5.2 Mechanisms that aid the effective retrieval of information from search engines

5.2.1 Geographical localisation

From an organisation's perspective, one way that search engines can help improve the relevance of search results is through geographical localisation. An organisation would benefit by having search results pertaining to regions where it has offices or business interests (or clients/customers). Sullivan (2003) explains how local commerce searching can often be a disappointing experience on the web's major search engines (a role performed well by the low-tech Yellow Pages). Major search engines are now joining the push for local search (a motivator being to tap on to local advertising revenue) and are allowing searches for local content, maps, driving directions, weather, *etc*. (Sullivan, 2003; Sterling, 2004; Mara, 2004; Notess, 2005).

5.2.2 Specialty search engines

While localisation refers to providing information more relevant to a geographical region (*e.g.*, Singapore or Eastern Europe), specialty search engines help provide information specific to an area or domain. An organisation would benefit from a search engine specialising in its industry of business, *e.g.*, a search engine each for people working on voice-over-IP, healthcare, biotech or digital cinema technology. If the big search engines are unable to deliver comprehensive access to the entire web, perhaps the time has come for more focused sites to offer near-comprehensiveness in their own chosen fields (Vortal, 2005; Khoussainov and Kushmerick, 2003; Battelle, 2004; Sullivan, 2000).

6 Usefulness of IPFF mechanisms to different searcher roles

Since a taxonomy provides a tree-like structure of documents/information in a repository, and helps users browse rather than search, it would be highly useful to the learner who is just starting to work on a new project and does not know exactly what he/she should be searching for. A location seeker who is simply trying to locate a document or piece of information will be able to browse iteratively through the nodes of a taxonomy or different branches until he/she is able to locate the item of the search. To the data gatherer and the focused searcher, the usefulness of a taxonomy may range from medium to high because they might be able to use the search feature effectively as well (as opposed to the browsing provided by a taxonomy), since they have a prior domain knowledge in the area of the search and might be able to phrase the search keyword effectively, increasing their chances of getting relevant results. However, this is

highly dependent on the implementation of the search engine or search service in use, and the user might still prefer to browse several levels into the taxonomy, until he/she gets the item of the search, or further refine the search keyword after browsing through the taxonomy.

As multiple views of the taxonomy and faceted classifications allow the same terms to be classified into different views, they would be highly useful to the learner in helping him/her choose the view of choice. The different views might also benefit the location seeker in locating a document faster, since there would be different paths to arrive at the same document or piece of data. The focused searcher will be able to quickly arrive at the answer to his/her question utilising a suitable view. The data gatherer would also benefit from multiple views though he/she might tend to use a particular view more than the others – one that more closely fits in with the prior domain knowledge of the data gatherer.

Personalisation would benefit all the four searcher roles as it would help the employee indicate to the search service that he/she is assuming the particular role, so that the service and results could be customised accordingly. As personalisation would take into account the prior domain knowledge of the employee, it would be highly useful to the data gatherer. Personalised service should also make the task of locating a document or piece of data easier for the location seeker, as the user's prior interaction with the service could have been logged, depending on the level and type of personalisation implemented. To the learner, who has no prior domain knowledge in the area of search, the level of usefulness of personalisation may range from high to medium depending on how well the level of personalisation succeeds in providing search results meant for the layman, and at the same time, keeps track of the learner's progress as he/she acquires more knowledge in the domain of search. Usefulness for the focused searcher may range from medium to high, depending on how well the search service determines the level/context of the question and the answers it is able to provide to the focused searcher.

Incorporating contextual cues from static/dynamic content/collaborative sources should benefit all the four roles of searchers by increasing the set of relevant results and decreasing the set of non-relevant results. Usefulness to the data gatherer may range from medium to high depending on the cues obtained from dynamic collaborative sources by matching the actions of the data gatherer with those of others with similar domain knowledge.

During a search, a learner or a location seeker may not use the right word or phrase to match what is understood by the search service provider. Thus, incorporation of semantics would greatly benefit these two searcher roles. In comparison to the learner or the location seeker, semantics would have lesser usage for the data gatherer and the focused searcher who have prior domain knowledge or a specific question and are able to form their queries more precisely.

A local search might have limited applicability for the learner or the data gatherer, unless the information they are searching for pertains to a geographical context. Localisation would be useful to the location seeker if he/she is searching for local information. It would be highly useful to the focused searcher seeking an answer to a specific locally relevant question, *e.g.*, the driving directions to a client site.

A specialty search would be extremely useful to the data gatherer as he/she would be able to access the search engine directly relevant to his/her prior domain knowledge and the domain of search. This would greatly benefit a focused searcher too, perhaps by

pulling out answers from the FAQ (Frequently Asked Questions) section of a specialised portal. It might have medium to high utility for the learner depending on whether the specialty search engine provides background or basic information that could be understood by a learner. Specialisation would have low utility for the location seeker, unless the location he/she is seeking resides within the specialty search engine.

Table 1 summarises the assumptions of usefulness of the different mechanisms to achieve a 'Fit' between information providers and information finders to the different roles searchers assume in an organisation. The usefulness is rated from low to medium to high, with high being the most relevant and useful. Experiments will be conducted and empirical evidence gathered to help us ascertain our assumptions of these ratings.

Table 1 Usefulness of different mechanisms to the different searcher roles

	Learner	*Data gatherer*	*Location seeker*	*Focused searcher*
Applicable to organisational knowledge repositories / search engines				
Taxonomy	High	High-medium	High	High-medium
Multiple views of the taxonomy	High	High-medium	High	High
Personalisation	High-medium	High	High	Medium-high
Usage of cues	High	Medium-high	High	High
Semantics	High	Medium	High	Medium
Applicable to search engines				
Localisation	Low-medium	Low-medium	Medium	High
Specialty search	Medium-high	High	Low	High

From the table, we can infer that a *learner* would benefit most from the usage of taxonomies, personalisation and the usage of cues/semantics. The *data gatherer* would benefit most from usage of taxonomies, personalisation and specialty search. The *location seeker* would benefit most from the usage of cues/semantics, as well as taxonomies and personalisation. The *focused searcher* would benefit immensely through the usage of cues, taxonomies and personalisation, as well as localisation and specialty search.

7 Case-study: Education Taxonomy Portal (ETaP)

We have developed an online portal '*Education Taxonomy Portal (ETaP)*'[3] as a case study to demonstrate the effectiveness of the seven attributes discussed above in enhancing search quality. While not directly applicable to an organisational perspective, it should nevertheless help us validate our assumptions of usefulness of the different mechanisms to achieve a 'Fit' between information providers and information finders to the different roles searchers assume.

ETaP is targeted towards the Singapore Education Community – primarily teachers and students who have to spend hours searching for relevant information on the internet, most of which is not directly relevant from a local Singapore perspective. It should also

help parents and all those associated with the education community in Singapore to perform quality search and be better satisfied with their search results. ETaP provides services to contribute, search, navigate and retrieve education-related content effectively.

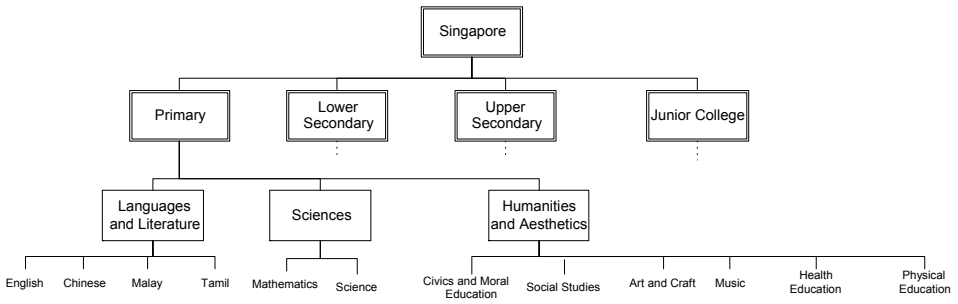
In the first phase, a portal for ETaP has been built. The next phase involves working with schools to gather content and to develop a pool of contributors. At the same time, the seven mechanisms described above will be built – some, in fact, are already in place. The final phase will involve developing and conducting experiments to validate the usefulness of mechanisms to the four searcher roles. The portal would then enter the maintenance phase, coupled with added contributions and continuous refinement. The model would subsequently be applied to schools in other regional countries.

In ETaP, we adopt the different mechanisms to achieve provider-finder fit for the four searcher roles as follows.

7.1 Taxonomy-based presentation and classification

A taxonomy based on the Singapore education curriculum facilitates ease of browsing. Information is organised based on the needs, perspective and vocabulary of primary and secondary schools and junior colleges in Singapore (see Figure 5). Taxonomies would later be developed based on the education curricula of other regional countries.

Figure 5 Section of ETaP taxonomy



7.2 Multiple views of the taxonomy

Multiple taxonomy views or faceted classification can be applied to ETaP in various ways. Apart from the syllabus view (Figure 5), there can be a student view, teacher view, knowledge area view, *etc*. Each view can be further classified based on date/freshness, subjects registered by student, *etc*. While a single taxonomy follows a top-down approach of classification (create a taxonomy and add documents/data to the nodes), multiple views would be facilitated by a bottom-up approach. Each document would have multiple keywords/tags associated with it, which would help determine its placement in different views and levels of the taxonomy. This approach would enable classification under different taxonomy views that are created on the fly, depending on the navigation pattern of the user during a particular browsing/searching session.

7.3 Personalisation

ETaP will have mechanisms for the searcher to specify whether he/she is a learner, a data gatherer, a location seeker or a focused searcher. For new users, the portal would also have help tools to enable users to be able to understand and determine their roles easily. The taxonomy will be dynamically modified, depending on the level and prior knowledge of the searcher. This prior knowledge will be captured through the level the student is studying at, and the average grade (an indicator of performance/knowledge level) of the student in the school. Teachers will also be directed to the level or class they are teaching. The built-in personalisation features will be further refined based on feedback from teachers and students and after a study of usage patterns over time.

7.4 Usage of cues

To further improve search quality, we intend to apply the different types of contextual cues in the system in ways described in Section 5.1.

7.5 Usage of semantics

Semantics will be built in to facilitate a better match of the user queries and the specialised terms used by creators of information.

7.6 Geographical localisation

Localisation is facilitated by narrowing the education-related content to the local Singapore context. Retrieved education content is current and conforms to the syllabus prescribed by the Singapore Education Council. As the local teachers and students are contributors (and moderators of information) as well as searchers, the relevance of content for the local education sector is significantly enhanced. Subsequently, the portal will be expanded to include countries and schools in the ASEAN region. Here, too, data will be kept relevant to the chosen countries.

7.7 Specialty search

ETaP provides specialisation by focusing on the Education domain (from the primary to secondary and junior college level). Tertiary education may be included based on a future feasibility study. Where relevant and depending on user needs, education-related material will also be gathered from major search engines and the search results combined.

Improved navigation and search quality provided by ETaP should give rise to more innovation and effectiveness, and enhance the efficacy of Knowledge Management in Singapore. Experiments will be conducted to measure the effectiveness of a search and to validate the hypotheses outlined in Section 2. Results will be reported in due course. ETaP is available free for everyone's use.

8 Conclusion

A global economy has emerged and become strong; and industrial economies and societies are being transformed into knowledge- and information-based service economies, where knowledge and information are key ingredients in creating wealth. Information retrieval forms a core part of knowledge management. There is a need to move beyond 'finding the needle in a haystack' to 'connecting the dots' among various pieces of information. In this quest to connect the dots, a 'one size fits all' model for information retrieval is inadequate. Three hypotheses have been formed. Knowledge repositories and search engines must cater to one of four roles searchers assume for information finding:

1 learner

2 data gatherer

3 location seeker

4 focused searcher.

Information providers and information finders must work hand-in-hand in order to achieve a 'Fit', leading to successful information finding by one of the four categories of roles of searchers. This Information Provider-Finder Fit can be achieved by combining a few or more of the following means: i) taxonomy-based presentation/classification; ii) multiple views of the taxonomy; iii) personalisation; vi) usage of cues; and vii) usage of semantics. A search from the internet would also benefit from vi) geographical localization; and vii) specialty-search engines. Experiments using ETaP will be conducted to try and prove/disprove these hypotheses. Improved search and navigation arising from the Fit should enhance the efficacy of Knowledge Management in an organisation and create information finders out of searchers of information.

References

Battelle, J. (2004) 'GlobalSpec: domain specific search and the semantic web', *John Battelle's Searchblog*, http://battellemedia.com/archives/000519.php

Cao, T.H., Nguyen, T.H.D. and Qui, T.C.T. (2003) *Searching the Web: A Semantics-Based Approach*, http://www.iwr.uni-heidelberg.de/HPSCHanoi2003/abstracts/cao_th_1.pdf

Clough, R. (2005) *Current Search Engine News*, K, Clough, Inc., viewed 29 March 2005, http://searchengineguide.com/searchenginenews.html

Daft, R. and Lengel, R. (1986) 'Organizational information requirements, media richness, and structural design', *Management Science*, Vol. 32, No. 5, pp.554–571.

Davenport, T.H. and Prusak, L. (2000) *Working Knowledge: How Organizations Manage What They Know*, Boston: Harvard Business School Press.

Davenport, T.H., DeLong, D.W. and Beers, M.C. (1998) 'Successful knowledge management projects', *Sloan Management Review*, Winter, Vol. 39, No. 2.

Davenport, T.H., Thomas, R.J. and Cantrell, S. (2002) 'The mysterious art and science of knowledge-worker performance', *MIT Sloan Management Review*, Fall, Vol. 44, No. 1.

Fleming, N.D. (1996) 'Coping with a revolution: will the internet change learning?', *Occasional Paper for Faculty*, Lincoln University, Canterbury, New Zealand, http://www.vark-learn.com /documents/Information_and_Knowle.pdf

Gao, F., Li, M. and Nakamori, Y. (2002) 'Systems thinking on knowledge and its management: systems methodology for knowledge management', *Journal of Knowledge Management*, Vol. 6, No.1, pp.7–17.

Garshol, L.M. (2004) 'Metadata? Thesauri? Taxonomies? Topic Maps! Making sense of it all', *Ontopia*, http://www.ontopia.net/topicmaps/materials/tm-vs-thesauri.html#N828

Goh, J.M., Poo, D.C.C. and Chang, K.T.T (2004) 'Incorporating contextual cues into electronic repositories', *Proceedings of the Eighth Pacific-Asia Conference on Information Systems*, Shanghai, China, 8–11 July, pp.472–484.

Gold, A.H., Malhotra, A. and Segars, A.H. (2001) 'Knowledge management: an organizational capabilities perspective', *Journal of Management Information Systems*, Summer, Vol. 18, No. 1.

Grover, V. and Davenport, T.H. (2001) 'General perspectives on knowledge management: fostering a research agenda', *Journal of Management Information Systems*, Summer, Vol. 18, No. 1.

Hendriks, P. and Virens, D. (1999) 'Knowledge-based systems and knowledge management: friends or foes?', *Information and Management*, Vol. 35, pp.113–125.

Khoussainov, R. and Kushmerick, N. (2003) 'Learning to compete in heterogenous web search environments', *Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence* (*IJCAI-03*).

Laudon, K.C. and Laudon, J.P. (2004) 'Management information systems: managing the digital firm', 8th (international) edition, *Pearson Education International*, Vol. 11, No. 314, pp.4–7.

Lawrence, S. and Giles, C.L. (1998) 'Searching the world wide web', *Science*, 3 April, Vol. 280, www.sciencemag.org

Loucopoulos, P. and Kavakli, V. (1999) 'Enterprise knowledge management and conceptual modelling', *Lectures Notes in Computer Science*, Vol. 1565, pp.123–143.

Mara, J. (2004) 'Local search: working hard for the money', *ClickZ News*, http://www.clickz.com/news/article.php/3333601

Martin, P. and Eklund, P. (1999) 'Embedding knowledge in web documents', *Proceedings of the 8th International World Wide Web Conference*.

Notess, G.R. (2004) 'Search engine news', *Search Engine Showdown*, http://searchengineshowdown.com/

Notess, G.R. (2005) 'Internet search engine update', *Online Magazine*, Information Today, Inc., November–December 2004, Vol. 28, No. 6, http://www.infotoday.com/online/nov04/SearchEngineUpdate.shtml

van Ossenbruggen, J. and Hardman, L. (2002) 'Smart style on the semantic web', *Semantic Web Workshop, WWW2002*, May, www.cwi.nl/~media/publications/INS-R0201.pdf

Papadopoullos, A. (2004) 'Answering the right questions about search, EContent leadership series – strategies for...search, taxonomy and classification', *Supplement to July/August 2004 EContent and Information Today*, pp.S6–S7, http://www.kmworld.com/publications/whitepapers/crm/EC_Search_04.pdf

Prusak, L. (1997) *Knowledge in Organizations*, Oxford: Butterworth-Heinemann.

Shapiro, Y. and Lehoczky, E. (2003) 'Search engine optimization', *SearchEngines.com*, http://www.searchengines.com/intro_optimize.html

Sterling, G. (2004) 'Local search: the hybrid future', *Search Engine Watch*, http://searchenginewatch.com/searchday/article.php/3296721

Sullivan, D. (2000) 'The Vortals are coming! The Vortals are coming!', *SearchEngineWatch*, http://searchenginewatch.com/sereport/article.php/2162541

Sullivan, D. (2003) 'Local search series', *SearchEngineWatch*, Parts 1–5, http://searchenginewatch.com/searchday/article.php/3111631

Sullivan, D. (2005) *Search Engine Watch*, Jupiter Media Corporation, viewed 29 March 2005, http://searchenginewatch.com/

Vortal, K. (2005) 'Why vortal', *Kawin Vortalbuilding.com*, Chicago: Kawin Interactive, Inc., http://www.vortalbuilding.com/whyvortal.html (viewed 29 March 2005).

Wigg, K.M. (1997) 'Knowledge management: where did it come from and where will it go?', *Expert System With Application*, Vol. 13, No. 1, pp.1–14.

Wilson, T. (2003) 'FacetMap: your home for faceted classification tools', *FacetMap*, Complete Information Architecture, Inc., 29 March 2005, http://facetmap.com

Zigurs, I. and Buckland, B.K. (1998) 'A theory of task/technology fit and group support systems effectiveness', *MIS Quarterly*, Vol. 22, No. 3, pp.313–334.

Notes

1 The traditional business firm was – and still is – a hierarchical, centralised, structured arrangement of specialists that typically relied on a fixed set of standard operating procedures to deliver a mass-produced product or service. The traditional management group relied – and still relies – on formal plans, a rigid division of labour, and formal rules (Laudon and Laudon, 2004).

2 Facets can be thought of as different axes along which documents can be classified (faceted classification), and each facet contains a number of terms. This would then describe the document from many different perspectives (Garshol, 2004). Taxonomy is a type of facet in which the headings are arranged into a hierarchy (Wilson, 2003).

3 Accessible from http://etap.comp.nus.edu.sg