An Analysis of Information Technology Taxonomy Literature towards Managing Knowledge

Junaiza Jarjis[a], Erne Suzila Kassim[b]

[a]*Kulliyyah of Information Communication and Technology*
International Islamic University Malaysia, 53100, Gombak, Kuala Lumpur
Tel : 03-61965601, Fax : 03-61965179
E-mail : junaizaj@gmail.com

[b]*Faculty of Office Management and Technology*
Universiti Teknologi Mara, 40450 Shah Alam, Selangor
Tel : 03-55435545, Fax : 03-55435576
E-mail : ernesu02@salam.uitm.edu.my

ABSTRACT

Taxonomy is a method used to ease information retrieval for an effective knowledge management system. Studies have found information technology is vital for knowledge to be successfully managed and shared. We decided therefore to survey IT taxonomy studies using literature review and classification of articles from 1989 to 2008 in order to explore how taxonomy applications have developed in this period. Based on 62 articles found, the study classifies the taxonomy research into main topics and scope, methodologies most often applied, theories and models, and contributing countries. Based on the findings, we offer suggestions on improving research in this field.

Keywords
Taxonomy, information technology, knowledge management

1.0 INTRODUCTION

Fleming (1996) points out a collection of information are not knowledge, but the pattern must be understood before it can be classified into knowledge (Bellinger, 2004). Yet, information overload always make the process of turning information into knowledge a complex issue. Thus, Sykes (2001) suggests a high quality indexing and a taxonomy design to ease information retrieval for effective knowledge management system. Taxonomy provides a map of the information available and it is a root aspect in knowledge management system. Its purpose is to ease information retrieval by offering consistent name scheme and up to date vocabulary (Jardine, 2002). Perceiving taxonomy as significance to knowledge management, it has started to gain a momentum in research (Cheung, Lee and Wang, 2005; Mane, 2006).

As part of knowledge management study, this paper focuses on the development of information technology taxonomy for managing knowledge through a review of related literature from 1989 to 2008. Information technology is found to be a key indicator to leverage knowledge (Raja and Erne, 2006).

Our main objective is to further seek answers for these questions: "What are the most researchable topics and areas in the IT taxonomy study?", "What are the methodologies commonly used?", "What are the relevant theories and models as the basis of the study?", and "Which country contributes the most to the field?"

2.0 DATA COLLECTION

Related articles were obtained from various online databases; Springer, Elsevier, ACM, IEEE, ProQuest, Emerald and Business Source Premier. We decided to include refereed journals and proceedings published by reputable research/ professional associations as our sources since they serve the significance of academic research work (Levy and Ellis, 2006). Keywords used for finding the articles were *taxonomy*, *knowledge taxonomy*, *taxonomy and knowledge management*, *taxonomy and conclusion* and *ontology*. A total of 67 articles, from the year 1989 to 2008, related to the keywords were obtained. Yet, only 62 were deemed useful for the research as 5 were either the content did not reflect the subject area, or they were not research papers. Table 1 and Table 2 illustrate the source of the articles and the year of publication respectively.

Table 1: Articles Analyzed

Article type	N	%
Refereed journal	27	43.55
Conference paper	35	56.45
Total	62	100.00

We found most of the articles (56.45%) were published in conference proceedings as compared to refereed journals. Hence, the lack of refereed, high quality, rigorous and valid research papers pose a question on the papers reputation and their contributions to the body of knowledge. Culnan (1978) argues although conferences are valuable scientific venues for exchange of ideas, and a

major incubator for new research agenda, the overall rigor of conference proceedings is lower than one found in leading journals (Levy and Ellis, 2006).

We also found majority of the articles (54.84%) were published and presented within the years 2001 and 2005. This may indicate an emergence of the trends in closely relating the importance of having classified characters of the items being revised. For instance, software quality and its development are being revised and classified into taxonomies for better evaluations (Blum, 1994; Mohamed, 2004; Wang, 2003). Programming languages are also given new categories in terms of approaches and algorithms (Ihantola, 2005; Kagdi, 2005; Karavirta, 2006; Murata, 2005). Information retrieval aspects (Chen, 2005; Viramachananeni, 2005) and network securities are also not left behind (Clincy, 2005; Mirkovic, 2004; Wang, 2003; Weaver, 2003).

Table 2: Year of Publication

Year	Refereed Journal (%)	Conference Paper (%)	Total (%)
Until 1995	5 : 8.06	2 : 3.23	7 : 11.29
1996 - 2000	3 : 4.84	3 : 4.84	6 : 9.68
2001 - 2005	15 : 24.19	19 : 30.65	34 : 54.84
2006 - Present	4 : 6.45	11 : 17.74	15 : 24.19
Total	27 : 43.55	35 : 56.45	62 : 100.00

The decreasing number of articles found from 2006 till present is not due to a declining interest but it is because of the obvious period difference. It has only articles published until January 2008. The 2 articles found in 2008 were showing trends of interest towards society (Gawron, 2008; Tang, 2008).

3.0 RESULTS

3.1 Topics and Scope

We used the SIGs from the ACM (Associated Computer Machinery) to classify the ICT subtopics. The topic for the *programming languages* is modified so that only one broad programming language class is used for this study. Thus, the other 2 SIGs that represent ADA and APL programming language are not included in this class. It is nearly impossible not to have more than 1 topic in each study found. However, this categorization is done due to the credence of the areas being researched upon. Table 3 illustrates the topics and the number of researches in the top five areas.

We found that most of the articles (16.13%) have attempted, created or discussed taxonomy in '*algorithm and computation theory*'. This may be because of the highly used method of algorithm in solving the problem of taxonomy creation. *Knowledge discovery in data* includes knowledge management processes, knowledge management systems and knowledge workers. Looking at it under the lights of knowledge management, we envisage it to be more of significance in the years to come. This is because even though the other researches are done in other areas, the purpose is the same, which is to classify the 'knowledge' they have in that area. Knowledge management being the core value of taxonomization is assumed to be the motivational factor for researchers to engage in it.

Table 3: Topics in IT Taxonomy

Topic	N	%
Algorithms and Computation Theory	10	16.13
Software Engineering	9	14.52
Security, Audit and Control	8	12.90
Data Communication	6	9.68
Knowledge Discovery In Data	6	9.68

Other topics identified are *computers and society* and *information retrieval* (6.45%); *management of data* (4.84%); *programming languages*, *computer architecture*, *artificial intelligence* and *computer-human interaction* (3.23%); *design automation*, *design of communication*, *electronic commerce*, and *hypertext, hypermedia and web* (1.61%). There are gaps still exist which no taxonomy study has been conducted for these topics*; accessible computing*, *applied computing*, *embedded systems*, *computer science education*, *genetic and evolutionary computation*, *computer graphics and interactive techniques*, *information technology education*, *measurement and evaluation*, *microarchitecture*, *management information systems*, *multimedia*, *mobility of systems users*, *data and computing*, *operating systems*, *symbolic and algebraic manipulation*, *simulation and modeling*, and *university and college computing services*.

3.2 Research Methodology

Our analysis of the method, which is divided into classification of theoretical studies and classification for empirical studies, drew on the study by Gonzalez, Gasco and Llopis (2006) with a minor revision to the empirical method. Field studies and field and case studies classifications are substituted with design and survey. Table 4 shows although both theory and empirical were evenly studied in the period of 1996 to 2000, the later has started to hit since then. This might illustrate the attempt of the studies to validate proposed models or to materialize the concepts. In the theoretical classification, applied concept studies were the most common with the main topics of discussions were *security, audit and control* and *information retrieval*, followed by conceptual and illustrative studies. On the other hand, design studies dominate the empirical classification research. Case study and survey however, did not provide a significant contribution to the empirical research. Most of the design studies focus on the topic of *algorithm and computation theory*.

3.3 Theories and Models

From the articles reviewed, we found none of the researcher has used any theory to develop the taxonomy. Using models as the foundation to the taxonomy design has only been applied widely since the early 2000. However, it is not common when only 17 (27.42%) of the studies have done so. Nonetheless, each researcher seems to use a different model. This, we assume is because of the diverse areas of study. Some of the models applied are knowledge model (Mohamed, 2004); cognitive model (Mosley, 2004); thesaurus model (Cheung, 2005); differencing content model for XML language (Murata, 2005); and dirichlet generative model (Viramachananeni, 2005). Lambe's (2007) knowledge lens framework and taxonomy work is newly developed. This could be as a guideline for future research on taxonomy development. Table 5 reveals the models used against the year.

Table 5: Models Utilized

Year	Refereed Journal (%)	Conference paper (%)	Total (%)
Until 1995	1 : 1.61		1 : 1.61
1996 - 2000	1 : 1.61	2 : 3.23	3 : 4.84
2001 - 2005	6 : 9.68	6 : 9.68	12 : 19.35
2006 - Present		1 : 1.61	1 : 1.61
Total	8 : 12.90	9 : 14.52	17 : 27.42

3.4 Contributing Countries

We believe it is imperative to analyze the countries that contribute to the taxonomy studies for understanding the trend better. We refer nationality to the university or other institutions for which the author(s) was working. We found 31 (50%) of the articles were from the United States of America. This may indicate taxonomy is a well accepted and studied area, and has a very strong presence among the scholars in the nation. Other countries were either from the European (27.42%) or the Asian continent (12.90%). There were also articles from South Africa and Australia, although their number is small.

4.0 CONCLUSION

Our study has shown the study of information technology taxonomy has progressively increased since the late 1980s. We conclude IT taxonomy studies tend to move towards empirical research. Although design study may be of interest to researchers, we suggest applying different social studies methodologies. The integration of qualitative and quantitative methods may produce more rigorous results and findings. Finally, we propose the models to be thoroughly built and tested for theory development as currently there is no theory available in the taxonomy for knowledge management study.

Table 4: Research Methodologies

	Until 1995 (%)	1996 - 2000 (%)	2001 - 2005 (%)	2006 - Present (%)	Total (%)
Total theoretical	5 : 8.06	3 : 4.84	14 : 22.58	5 : 8.06	27 : 43.55
Conceptual	2 : 3.23	1 : 1.61	7 : 11.29	1 : 1.61	11 : 17.74
Illustrative	1 : 1.61	1 : 1.61	1 : 1.61		3 : 4.84
Applied-concept	2 : 3.23	1 : 1.61	6 : 9.68	4 : 6.45	13 : 20.97
Total empirical	2 : 3.23	3 : 4.84	20 : 32.26	10 : 16.13	35 : 56.45
Design	1 : 1.61	3 : 4.84	16 : 25.81	8 : 12.90	28 : 45.16
Case study			3 : 4.84	1 : 1.61	4 : 6.45
Survey	1 : 1.61		1 : 1.61	1 : 1.61	3 : 4.84
Total	7 : 11.29	6 : 9.68	34 : 54.84	15 : 24.19	62 : 100.00

REFERENCES

Bellinger, G. (2004). Knowledge management —emerging perspectives. Retrieved Feb 25, 2008 from http://www.systems-thinking.org/kmgmt/ kmgmt.htm

Blum, B. I. (1994). A taxonomy of software development methods. *Communications of the ACM, 37*(11), 82-94.

Bradley, E. H., Curry, L.A., and Devers, K.J. (2007). Qualitative data analysis for health services research: developing taxonomy, themes and theory. *Health Services Research, 42*(4), 1758-1772.

Chen, Y., Bart Jr., H.L., and Teng, F. (2005, November 10-11). *A content-based image retrieval system for fish taxonomy* in MIR '05, Singapore.

Cheng, C. P., Lau, G.T., and Law, K.H. (2007, June 4-8). *Mapping regulations to industry-specific taxonomies* in ICAIL '07, Palo Alto, California, USA.

Cheng, P. J., Tsai, C.H., Hung, C.M., and Chien, L.F. (2006). *Query taxonomy generation for web search* in CIKM'06, Arlington, Virginia, USA.

Cheung, C. F., Lee, W.B., and Wang, Y. (2005). A multi-facet taxonomy systems with applications in unstructured knowledge management. *Journal of Knowledge Management, 9*(6), 76-91.

Chignell, M. H. (1990). A taxonomy of user interface terminology. *SIGCHI Bulletin, 21*(4).

Chung, C. Y., Lieu, R., Liu, J., Luk, A., Mao, J., and Raghavan, P. (2002). *Thematic mapping–from unstructured documents to taxonomies* in CIKM '02, USA.

Clincy, V. A., and Abu-Halaweh, N. (2005). *A Taxonomy of free network sniffers for teaching and research* in Consortium for Computing Sciences in Colleges, Midwestern Conference.

D?zeroski, S. (2007). *Towards a general framework for data mining* in KDID'06, LNCS 4747.

Darvill, T. (2002). *The concise oxford dictionary of archaeology*. Retrieved February,20 2008 from http://www.oxford reference.com/views/ENTRY.html?subview=Main &entry=t102.e4170

Denning, D. E., and Branstad, D.K. (1996). A taxonomy for key escrow encryption system. *Communications of the ACM, 39*(3), 34-40.

Dombroviak, K. M., and Ramnath, R. (2007). *A taxonomy of mobile and pervasive applications* in SAC '07, Seoul, Korea.

Fishkin, K. P. (2004). A taxonomy for and analysis of tangible interfaces. *Pers Ubiquit Comput, 8*, 347-358.

Gawron, V. J., Drury, C.G., Czaja, S.J., Wilkins, D.M. (2008). A taxonomy of independent variables affecting human performance. *International Journal of Man-Machine Studies*, 643-671.

Gonzalez, R., Gasco, J., and Llopis, J. (2006). Information systems outsourcing: A literature analysis. *Information & Management, Vol 43*, 821–834.

Grimshaw, D. J. (1996). *Towards a taxonomy of geographical information systems* in Proceedings of the 29th Annual Hawaii International Conference on System Sciences, Hawaii.

Grover, V., and Goslar, M. (1993). Toward an empirical taxonomy and model of evolution for telecommunication technologies. *Journal of Information Technology*, 167-176.

Hahn, U., and Mark´o, K.G. (2001). *Joint knowledge capture for grammars and ontologies* in K-CAP'01, Victoria, British Columbia, Canada.

Hu, Y., Zheng, Q., Bai, H., Sun, X., and Dang, H. (2005). *Taxonomy Building and Machine Learning Based Automatic Classification for Knowledge-Oriented Chinese Questions* in ICIC 2005, Part I, LNCS 3644.

Ihantola, P., Karavirta, V., Korhonen, A., and Nikander, J. (2005). *Taxonomy of effortless creation of algorithm visualizations* in ICER'05, Seattle, Washington, USA.

Jardine, D. (2002). User-centered Taxonomies, *Sydney Knowledge Management Forum.*

Jayatilaka, B., and Lee, J. (2003). *An epistemological taxonomy for knowledge management systems analysis* in Proceedings of the 36th Hawaii International Conference on System Sciences, Hawaii.

Jonsson, P. (2000). An empirical taxonomy of advance manufacturing technology. *International Journal of Operations and Production Management, 20*(12), 1446-1474.

Kagdi, H., Collard, M.L., and Maletic, J.I. (2005, May 17, 2005). *Towards a taxonomy of approaches for mining of source code repositories* in MSR '05, Saint Louis, Missouri, USA.

Kakabadse, N. K., Kakabadse, A., and Kouzmin, A. (2003). Reviewing the knowledge management literature: Towards a taxonomy. *Journal of Knowledge Management, 7*(4), 75- 91.

Karavirta, V., Korhonen, A., and Malmi, L. (2006, Sept. 4-5). *Taxonomy of algorithm animation languages* in SOFTVIS '06, Brighton, UK.

Kerschberg, L., Baum, R., Waisanen, A., Huang, I. and Yoon, J. (1991). *Managing faults in telecommunications networks: A taxonomy to knowledge-based approaches* in Proceedings of IEEE International Conference on Systems, Man and Cybernetics, Man and Cybernetics.

Kouadio, M., and Pooch,U. (2002). A taxonomy and design considerations for internet accounting. *ACM SIGCOMM Computer Communications Review, 32*(5), 39-48.

Lambe, P. (2007). *Organizing knowledge: Taxonomies, knowledge and organizational effectiveness*. Oxford, England: Chandos Publishing.

Levy, Y. E., T. J. (2006). A systems approach to conduct an effective literature review in support of information systems research. *Informing Science Journal, 9*, 182 - 212.

Lytras, M. D., and Pouloudi, A. (2006). Towards a development of a novel taxonomy of knowledge management systems from a learning perspective: An integrated approach to learning and knowledge infrastructures. *Journal of Knowledge Management, 10*(6), 64-80.

Mehta, N. R., Medvidovic, N., and Phadke, S. (2000). *Towards a taxonomy of software connectors* in ICSE' 00, Limerick, Ireland.

Mirkovic, J., and Reiher, P. (2004). A taxonomy of DDOS attack and DDOS defense mechanisms. *SIGCOMM Computer Communications Review, 34*(2), 39-54.

Modahl, M., Agarwalla, B., Saponas, T.S., Abowd, G., and Ramachandran, U. (2006). UbiqStack: a taxonomy for a ubiquitous computing software stack. *Pers Ubiquit Comput, 10*, 21-27.

Mohamed, A. H., Lee, S.P., and Salim, S.S. (2004). An ontology-based knowledge model for software experience management. *Journal of Knowledge Management Practice*.

Möller, S. (2002, July). *A new taxonomy for the quality of telephone services based on spoken dialogue systems* in Proceedings of The Third SIGdial Workshop on Discourse and Dialogue, Philadelphia.

Mosley, P. (2004). *A taxonomy for learning object technology* in Consortium for Computing Sciences in Colleges, Eastern Conference.

Murata, M., Lee, D., Mani, M., and Kawaguchi, K. (2005). Taxonomy of XML schema languages using formal language theory. *ACM Transactions on Internet Technology, 5*(4), 660-704.

Nahrstedt, K., and Balke, W.T. (2004, October 10-16). *A Taxonomy for multimedia service composition* in MM'04, USA.

Ngassam, E. K., Kourie, D.G., and Watson, B. (2006). *A taxonomy of DFA-based string processors* in Proceedings of Saicsit 2006.

Olivier, M. S., and Von Solms, S.H. (1994). A taxonomy for object-oriented secure databases. *ACM Transactions on Database Systems, 19*(1).

Oman, P. W., and Cook, C.R. . (1990). *A taxanomy for programming style* in Proceedings of the 1990 ACM Annual Conference on Cooperation.

Pérez, M. P., Sánchez, A.M., de Luis Carnicer, M.P., and Vela Fiminéz, M.J. (2002). Knowledge tasks and teleworking: A taxonomy model of feasibility adoption. *Journal of Knowledge Management, 6*(3), 272-284.

Qi, Y., Candan, K.S., Sapino, M.L., and Kintigh, K.W. (2007, June 12-14). *Integrating and querying taxonomies with quest in the presence of conflicts* in SIGMOD '07, Beijing, China.

Raja Kasim, R. S., and Kassim, E.S. (2006). Knowledge management practices among MSC status organizations: a survey. *International Journal of Knowledge, Culture and Change in Organizations, Vol. 5*.

Ramamritham, K., and Chrysanthis, P.K. (1996). A taxonomy of correctness criteria in database application. *The VLDB Journal, 5*, 85-97.

Ramirez, Y. W., and Nembhard, D.A. (2004). Measuring knowledge worker productivity: A taxonomy. *Journal of Intellectual Capital, 5*(4), 602-628.

Saeed, H., and Chaudhry, A.S. (2002). Using Dewey decimal classification scheme (DDC) for building taxonomies for knowledge organisation. *Journal of Documentation, 58*(5), 575-583.

Savola, R. M. (2007, October 29). *Towards a taxonomy for information security metrics* in QoP '07, Virginia, USA.

Sawyer, S., Fedorowicz, J., Tyworth, M., Markus, M.L., and Williams, C.B. (2007). *A taxonomy for public safety* in Proceedings of the 8th Annual International Digital Government Research Conference.

Shoniregun, C. A. (2004). *Classification and taxonomy of TEISMEs* in ICEC '04, Sixth International Conference on Electronic Commerce, Delft, The Netherlands.

Spangler, S., and Kreulen, J. (2002, November 4-9). *Interactive methods for taxonomy editing and validation* in CIKM'02, McLean, Virginia, USA.

Spangler, S., Kreulen, J.T., and Lessler, J. (2003). Generating and browsing multiple taxonomies over a document collection. *Journal of Management Information System, 19*(4), 191-212.

Succi, G., Baruchelli, F., and Ronchetti, M. (1996). *A taxonomy for identifying a software component from uncertain and partial specifications* in Symposium on Applied Computing.

Sykes, J. (2001). *The value of indexing*. Retrieved December 16, 2007 from http://www.factiva.com/infopro/indexingwhitepaper.pdf

Tang, L., Liu, H., Zhang, J., Agarwal, M.N., and Salerno, J.J. (2008). Topic taxonomy adaptation for group profiling. *ACM Transactions on Knowledge Discovery from Data, 1*(4), 15.10-15.28.

Tang, L., Zhang, J., and Liu, H. (2006, August 20-23). *Acclimatizing taxonomic semantics for hierarchical content classification* in KDD'06, Philadelphia, Pennsylvania, USA.

Tilak, S., Abu-Ghazaleh, N.B., and Heinzelman, W. (2002). A taxonomy of wireless micro-sensor network models. *Mobile Computing and Communications Review, 6*(2).

Tseng, M. C., Lin, W.Y., and Jeng, R. (2005). *Maintenance of generalized association rules under transaction update and taxonomy evolution* in DaWaK 2005, LNCS 3589, Berlin, Heidelberg.

Tzitzikas1, Y., Spyratos, N., and Constantopoulos, P. (2005). Mediators over taxonomy-based information sources. *VLDB Journal, 14*, 112-136.

Valiati, E. R. A., Pimenta, M.S., and Freitas, C.M.D.S. (2006). *A taxonomy of tasks for guiding the evaluation*

of multidimensional visualizations in BELIV 2006, Venice, Italy.

Vanthournout, K., Deconinck, G., and Belmans, R. (2005). A taxonomy for resource discovery. *Pers Ubiquit Comput, 9*, 81-89.

Venugopal, S., Buyya, R. and Ramamohanarao, K. (2006). A taxonomy of data grids for distributed data sharing, management and processing. *ACM Computing Surveys, 38*.

Viramachananeni, S., Sona, D., and Avosani, P. (2005). *Hierarical dirichlet for document classification* in Proceedings of the 22nd Int Conf on Machine Learning Bonn, Germany.

Vitolo, T. M., and Coulston, C. (2002). Taxonomy of information literacy competencies. *Journal of Information Technology Education, 1*(1).

Wang, H., and Wang, C. (2003). Taxonomy of security considerations and software quality. *Communications of the ACM*, 75-78.

Weaver, N., Paxson, V., Staniford, S., and Cunningham, R. (2003, October 27). *A taxonomy of computer worms* in WORM '03, Washington DC, USA.

Wemmerlöv, U. (1989). A taxonomy for service processes and its implications for system design. *IJSIM, 1*(3), 20-40.

Whitman, L., Ramachandran, K., and Ketkar, K. (2001). *A taxonomy of living model of the enterprise* in Proceedings of the 2001 Winter Simulation Conference.

Widdows, D. (2003, May-June 2003). *Unsupervised methods for developing taxonomies by combining syntactic and statistical information* in Proceedings of HLT-NAACL 2003, Main Papers, Edmonton.

Wun, A., Cheung, A., and Jacobsen, H.A. (2007, June 20-22). *A taxonomy for denial of service attacks in content-based publish/subscribe* in DEBS '07, Toronto, Ontario, Canada.

Yehezkel, C. (2002, June 24-26). *A taxonomy of computer architecture visualization* in ITiCSE '02, Aarhus, Denmark.

Zeimpekis, V., Giaglis, G.M. and Lekakos, G. (2003). A taxonomy of indoor and outdoor positioning techniques for mobile location services. *ACM SIGECOM Exchanges, 3*(4), 19-27.

Ziegler, C. N., Lausen, G., and Thieme, L.S. (2004, November 8-13). *Taxonomy-driven computation of product recommendations* in CIKM'04, Washington, D.C., USA.