Building an IT Taxonomy with Co-occurrence Analysis, Hierarchical Clustering, and Multidimensional Scaling

Chia-jung Tsui, Ping Wang, Kenneth R. Fleischmann, Asad B. Sayeed, and Amy Weinberg
University of Maryland, College Park, MD 20742
{ctsui, pwang, kfleisch, asayeed, weinber}@umd.edu

ABSTRACT

Different information technologies (ITs) are related in complex ways. How can the relationships among a large number of ITs be described and analyzed in a representative, dynamic, and scalable way? In this study, we employed co-occurrence analysis to explore the relationships among 50 information technologies discussed in six magazines over ten years (1998-2007). Using hierarchical clustering and multidimensional scaling, we have found that the similarities of the technologies can be depicted in hierarchies and two-dimensional plots, and that similar technologies can be classified into meaningful categories. The results imply reasonable validity of our approach for understanding technology relationships and building an IT taxonomy. The methodology that we offer not only helps IT practitioners and researchers make sense of numerous technologies in the iField but also bridges two related but thus far largely separate research streams in iSchools – information management and IT management.

Keywords

Information technology management, taxonomy, co-occurrence, hierarchical clustering, multidimensional scaling

1. INTRODUCTION

The proliferation of information technologies (ITs) has been accompanied by the proliferation of information in recent decades. Opportunities emerge from such proliferation of information and technologies, making the iField an increasingly prominent and vibrant area for research and practice. At the same time, just as the explosion of information presents serious challenges in information management, the seemingly everlasting propagation of numerous ITs poses challenges in IT management. The bewildering amount of IT confronting IT practitioners and researchers renders it a challenging task to make sense of the technologies, in order to effectively manage or productively study them. In practice, IT management has been traditionally undertaken along functional lines such as hardware, software, networking, and services. Streams in IT management research, on the other hand, have mapped well onto traditional categories in practice, drawing insights from various reference disciplines such as computer science, psychology, economics, and sociology. However, recent technological and managerial advances have blurred the boundaries of traditional categories. For example, software and service have converged under the rubric of "software as a service" (SaaS). Moreover, because different types of IT may entail different cost structures, work processes, and potential returns, different ITs may require different management practices and different research methodologies. Hence, contemporary IT management practices (such as IT portfolio management) and the increasing emphasis on interdisciplinary research call for rigorous and up-to-date classifications, or *taxonomies*, of IT.

Thus far, while it has been argued that various ITs are related to varying degrees [26], it is still difficult to make sense of the relationships among technologies. For example, here is a partial list of contemporary ITs: service-oriented architecture (SOA), Web services, open source software (OSS), Web 2.0, YouTube, iPhone, blogs, and cloud computing. How are they related? How can we measure their similarities and differences? How might they be classified into meaningful categories?

IT practitioners and researchers are not well equipped to answer these questions. On the one hand, many studies in the dominant paradigm of IT management research have demonstrated that various organizational, technical, and environmental factors influence IT adoption and use [10]. As this dominant paradigm is reaching "the point of diminishing returns as a framework for supporting ground-breaking research" [10, p. 314], we note that most studies in the dominant paradigm employ single-technology research designs, leaving the relationships among ITs under-explored. On the other hand, a peripheral, yet sustained stream in IT management research has employed a variety of methods to classify technologies, practices, and/or research topics in IT [see a most recent review in 6]. Most studies in this stream had to explicitly or implicitly rely on domain experts to evaluate the similarities or differences among various technologies [e.g., 7]. Although experts may be skillful in detecting the subtleties within and across the different types of technologies, expert evaluation is often (1) biased towards the views of specific experts contributing to specific studies, (2) static in the time of such evaluations, and (3) difficult to scale up to examine the relationships among a large number of ITs. Therefore, considering the current status of the IT management literature, we raise this research question: *How can the relationships among a large number of ITs be described and analyzed in a representative, dynamic, and scalable way?*

We address this question in this study by offering a representative, dynamic, and scalable methodology to understand technology relationships and build IT taxonomies. Our approach, combining co-occurrence analysis, hierarchical clustering, and multidimensional scaling, lends itself well to automation and complements extant expert-based methods. In the following, we first briefly review the current approaches to taxonomy creation in general and IT taxonomy specifically. Then we illustrate our approach with an empirical study of 50 ITs over ten years. And finally we conclude by discussing the validity, benefits, and limitations of our approach for IT management research and practice.

2. TAXONOMY AND IT TAXONOMY

2.1 Taxonomy for Information Management

Today organizations are employing an increasing number of types of IT and an increasing amount of IT. Consequently, the amount of data collected and stored by various ITs increases exponentially. Content services, a popular technique in information management, migrate data from various sources to a common pool. Because content services do not provide inherent organizational structure for the pooled data, a taxonomy should be created to allow users to efficiently and effectively browse and use information. Hence, taxonomy creation is an important element of content management in organizations [14, 16].

A taxonomy is a classification scheme (often hierarchical) of information components (for example, terms, concepts, graphics, sounds) and their interrelationships [13]. Taxonomy creation is usually a "top-down process" by which domain experts provide an overview of the domain, list categories and features of each category, and finally classify categories into broader classes according to how similar the features of the categories are [17]. Categories that do not mach current classes are put aside until enough categories with sufficiently similar features appear to justify the creation of new classes [16]. It has been recommended that analysts use and customize pre-populated taxonomies whenever available [14].

2.2 IT Taxonomy

Sustained, despite relatively peripheral, efforts have been made in IT research to classify technologies, applications, and research topics and methods [e.g., 1, 2, 6, 7, 9, 11, 23, 25]. The usual process that produces various classification schemes and taxonomies in IT is very similar to the taxonomy creation approach for information/content management described above, except that the initial list of categories that constitute the taxonomies often comes from empirical surveys of the IT discourse. For example, three classification studies of information systems research collected keywords of publications as initial input to the taxonomy creation process [1, 2, 6]. Once the initial list has been compiled, experts (often the study's authors) arrange the items on the list according to their assessment of various features of the items. For example, Ein-Dor and Segev [7] surveyed the definitions of 17 technologies in the IT discourse, identified from the definitions 31 attributes and 27 functions, and then described the technologies by two bit-vectors: a vector of attributes and a vector of functions. Furthermore, they performed quantitative methods such as multidimensional scaling (MDS) to visualize the relationships among the technologies in terms of their relative similarity.

2.3 Limitations of Extant Approaches

To varying degrees, extant methods for creating taxonomies in general and specifically for IT rely on experts. While expert opinions are valuable in grounding the taxonomy in specific domains and detecting subtleties in the relationships among categories, current approaches have several limitations.

First, the structures of extant taxonomies represent a relatively narrow set of views from only a few experts. For instance, the choice of features (such as attributes and functions of IT) for classification depends on the specific opinions or background knowledge of the experts who participate in the study. Second, taxonomies built by this approach seem static, fixed at the time when experts created them. Efforts to update existing taxonomies are few and far in between. For example, the ACM Computing Classification System currently being used was created in 1998. As another example, the official Keyword Classification Scheme for Information System Research was last updated in 1993 [2]. Finally, such scant efforts to update existing taxonomies may be due to another limitation – methods relying on experts are not scalable, lending themselves poorly to automation. As the number of ITs increases, the effort by human experts to describe each technology according to its attributes and functions increases, and the reliability of that classification work may decrease. Addressing these limitations of the extant approaches to IT taxonomy creation, in this study we develop a methodology that builds upon the existing methods of analyzing IT discourse, but allows wider representations of opinions, dynamic updating at multiple points of times, and large-scale automated analysis of a large number of technologies.

3. DEVELOPING A REPRESENTATIVE, DYNAMIC, AND SCALABLE APPROACH

In this section, we describe our approach with an empirical study as an illustration.

3.1 Data Collection

There are many outlets for the IT discourse, including books, magazines, conferences, blogs, wikis, and many others, where discourse data may be collected. In order to illustrate how our methodology works in IT management research, we decided to focus on two IT trade magazines (*InformationWeek* and *Computerworld*), two business magazines (*BusinessWeek* and *The Economist*), and two news magazines (*Newsweek* and *US News & World Report*). As described below, the scale of the data we collected from the six magazines is large enough for us to demonstrate the scalability of our approach. In addition to the scale, our data is also diverse, representing a wide range of views on IT and broader topics.

We downloaded from the Lexis/Nexis online database all articles published during a ten-year period (1998-2007) in the six magazines, totaling about 220,000 articles. Meanwhile, we compiled a list of 50 IT concepts (Table 1), ranging from enterprise software (e.g., CRM) to personal gadgets (e.g., iPod), from abstract concepts (e.g., artificial intelligence) to concrete products/services (e.g., YouTube), and from highly popular (e.g., e-business) to less well-known concepts (e.g., digital subscriber line – DSL). Admittedly, this list is *ad hoc*, but it serves the illustration purpose because the list covers a broad range of technologies in the examination period. We then extracted from the articles all paragraphs that contain any of the technologies on the list. In doing so, we considered multiple possible labels for each technologies, plural forms, and acronyms unique to the technology. For example, in extracting paragraphs containing "digital subscriber line," we also included paragraphs mentioning "digital subscriber lines" and "DSL." In total, 105,400 paragraphs containing at least one technology on the list were extracted from the full text of the articles published the six magazines.

Table 1. Information technologies examined in the study

Label	Full Name of Information Technologies
AI	Artificial intelligence
ASP	Application service provider
BI	Business intelligence
Blog	Blog
Bluetooth	Bluetooth
BizProReen	Business process reengineering
CloudCom	Cloud computing
CRM	Customer relationship management
DigiCam	Digital camera
DLearn	Distance learning
DSL	Digital subscriber line
DecisionSS	Decision support system
DW	Data warehouse
eBiz	Electronic business
eCom	Electronic commerce
EDI	Electronic data interchange
ERP	Enterprise resource planning
ExpertSys	Expert system
GPS	Global positioning system
Grpware	Groupware
IM	Instant messaging
iPhone	iPhone
iPod	iPod
KM	Knowledge management
Linux	Linux
Multimedia	Multimedia
MP3	MP3 player
MySpace	MySpace
NeuralNet	Neural net
OLAP	Online analytical processing
OSS	Open source software
Outsource	Outsourcing
PDA	Personal digital assistant
RFID	Radio frequency identification
SmartCard	Smart card
SCM	Supply chain management
SFA	Salesforce automation
SocNet	Social networking
SOA	Service oriented architecture
Telecommute	Telecommuting
TabletPC	Tablet PC
UtiComp	Utility computing
Virtualization	Virtualization
VPN	Virtual private network
Web2	Web 2.0
WebServ	Web services
WiFi	Wi-Fi
Wiki	Wiki
Wikipedia	Wikipedia
YouTube	YouTube

3.2 Data Analysis

To make sense of the relationships among the technologies, we focused on the initial step of exploring the similarity of the technologies. One approach is to automatically infer similarity of technologies from their co-occurrences in the same unit of discourse (e.g., an article, paragraph, or sentence) [22]. We also used hierarchical clustering analysis and multidimensional scaling to classify the technologies.

3.2.1 Co-occurrence

Co-occurrence of words or terms has been used in various fields such as computational linguistics [3, 4] and information retrieval [21] to study the relationships among words or terms. For example, Spence and Owens [22] used co-occurrence to evaluate the strength of word association. They found that related pairs of nouns co-occur considerably more often than unrelated pairs. Their finding suggests that co-occurrence frequency may indicate the strength of word association.

Analysis of co-occurrence should define a proper size of the window where words or terms co-occur. A window size can be a certain number of words or characters [e.g., a window of 250 characters in 22] or a logical division of an input text [19]. We chose paragraph as the window size because it sufficiently captures the context for describing related technologies.

To measure co-occurrence at the paragraph level, from the 105,400 paragraphs we initially extracted, we selected paragraphs containing two or more ITs in Table 1. This filtering process returned approximately 12,000 paragraphs. Then we constructed a 50x50 co-occurrence matrix with each row or column representing a technology on the list. The value in each cell of the matrix represents the number of paragraphs containing the respective pair of technologies. This co-occurrence matrix is a matrix of similarity. In order to perform subsequent classification and visualization techniques that are based on dissimilarity measures, we transformed the similarity matrix to a dissimilarity matrix with the formula: *1/(x+0.1)*.

3.2.2 Hierarchical clustering

Cluster analysis is the process of grouping objects into unknown clusters such that the within-group variation is minimized and the between-group variation maximized [8]. The agglomerative hierarchical clustering method groups objects on a series of levels, from the finest partition, in which each individual object forms its own cluster, and successively combines smaller clusters into larger ones until all objects are in one cluster. Agglomerative hierarchical clustering employs an aggregation criterion, or "linkage rule," to determine how the distance between two clusters should be calculated based on the distance scores of pairs of objects. The most well-known aggregation criteria are single link, complete link, and average link [12]. The distance between two clusters is represented by the minimum, maximum, or average distance between any pair of objects, one object from each cluster. In single link clustering, two clusters with the smallest minimum pairwise distance are merged in each step. In complete link clustering, two clusters with the smallest maximum pairwise distance are merged in each step. And average link clustering is a compromise between the other two methods. We used the average link in this study because of its robustness [5].

3.2.3 Multidimensional scaling

Previous research has found that applying multidimensional scaling (MDS) and clustering separately to the same proximity data results in greater insight into the structure underlying the data and can detect more subtle and complex relationships than either method used alone [15, 18, 20]. Both clustering and MDS are visualization techniques. The key difference between the two techniques is that MDS provides a spatial representation of the data, while clustering provides a tree representation [15].

Based upon a matrix of item-item similarities or dissimilarities, an MDS algorithm assigns a location to each item in a space such that the distances between the items correspond as closely as possible to the measured dissimilarities between the items. In other words, the proximity of items to each other in the space indicates how similar they are. We used the MDS procedure based on the ALSCAL or alternating least squares scaling [24], a popular MDS algorithm. For easy interpretation of the result, we chose to present the MDS solutions in two-dimensional scatter plots.

4. RESULTS

Our clustering analysis of the transformed co-occurrence matrix generated a hierarchical structure of 50 technologies in a dendrogram (Figure 1), where vertical lines show joined clusters and the position of the lines on the scale from 1 to 25 indicates the distance at which clusters are merged. By inspecting the dendrogram, we have identified eight clusters, all of which merged around 5 in the 25-point scale. These eight clusters are indicated by the intersections between the dendrogram and the vertical dotted line in Figure 1. Table 2 summarizes the membership of each cluster. In Figure 2, we depict the 50 ITs in a two-dimensional MDS plot. Following Shepard and Arabie's [20] suggestion, we have used different colors to represent the eight clusters identified in the clustering analysis. Generally speaking, most of the technologies in the same cluster are located close to each other in the MDS plot. We describe several clusters in more details below.

Table 2. Membership of the clusters

Cluster	Labels of Information Technologies*
1	eBiz, eCom, CRM, ERP, Outsource, ASP, SCM, SFA, EDI, Grpware, KM, BizProReen (BPR)
2	RFID, SmartCard
3	BI, DW, OLAP, DecisionSS
4	AI, NeuralNet, ExpertSys
5	DSL, VPN, Telecommute, DLearn
6	Bluetooth, WiFi, PDA, GPS, iPod, MP3, DigiCam, Multimedia, iPhone, TabletPC
7	Wiki, Wikipedia, MySpace, SocNet, Blog, YouTube, Web2.0, IM
8	UtiComp, Virtualization, Linux, OSS, SOA, WebServ, CloudCom

* Please see the full names of the IT labels in Table 1.

Cluster 1 includes twelve IT concepts. All of them are enterprise IT applications except outsourcing, which is a strategy for managing enterprise IT. Business process reengineering (BPR) was the last to join the cluster, suggesting that it is the least similar to the others in the cluster. This situation may explain why BPR looks like an outlier in the cluster in the MDS plot (Figure 2). Cluster 5 includes four IT concepts. Among them, digital subscriber line and virtual private network are both telecommunication technologies, which may be employed in the other two IT applications (telecommuting and distance learning). Cluster 6 has ten IT concepts, all related to mobile or wireless technologies. Some, such as bluetooth and Wi-Fi, are the underlying mobile technologies. Others, such as TabletPC and PDA, are the devices enabled by the wireless/mobile technologies. Cluster 7 has eight IT concepts. They are the so-called Web 2.0 technologies that have become highly popular in recent years. Lastly, Cluster 8 includes seven IT concepts of similar type such as utility computing, Web service, and cloud computing.

According to the agglomeration schedule, a series of steps during clustering, we were able to identify twelve pairs of ITs considered most similar to each other in the list (Table 3). The pairs include, for example, e-business and e-commerce, iPod and MP3, and artificial intelligence and neural net. These pairs are compatible with even rudimentary understanding of these technologies.

Table 3. Pairs of most similar ITs

Pair	IT*	Pair	IT*
1	eBiz, eCom	7	Bluetooth, WiFi
2	CRM, ERP	8	iPod, MP3
3	Linux, OSS	9	DSL, VPN
4	BI, DW	10	Grpware, KM
5	SOA, WebServ	11	AI, NeuralNet

6	MySpace, SocNet	12	Wiki, Wikipedia

* Please see the full names of the IT labels in Table 1.

5. DISCUSSION

5.1 Validity of the Approach

The results illustrate that co-occurrence data can be utilized for classification. Our co-occurrence analysis, supplemented by the two visualization techniques, has yielded results that can be interpreted fairly easily, even without the presence of sophisticated expert knowledge of the various domains that our list covers. The face validity we have seen in this illustration study gives us reasonable confidence in applying our methodology to other circumstances where *a priori* knowledge is unavailable, such as the cases of new or unknown technologies.

5.2 Benefits of the Approach

Our approach has several advantages. First, this approach is more representative than extant methods for taxonomy creation, which often rely on a small number of experts and represent a narrow set of views. The community of organizational and individual stakeholders represented by any of the magazines we selected to study in this project is obviously larger than any group of experts enlisted in previous classification studies. Out of curiosity, we sorted the data by each magazine and performed the same analysis. Figure 3 compares the dendrograms we produced using the *InformationWeek* and *BusinessWeek* data. The differences in the IT hierarchies signify the different structures of IT knowledge that were developed in the two communities. In the illustration above, we pooled the data from all six magazines, making the results even more representative of the broader socio-technical context in which technological innovations emerge and evolve.

Second, speaking of evolution, we note that our approach allows updating taxonomies at multiple points of time, enabling longitudinal analysis of the dynamic relationships among technologies. In fact, technologies do change over time and their relationships change too. For illustration, we divided the *InformationWeek* data into two five-year periods (1998-2002 and 2003-2007) and performed the same analysis on the two subsets of data. Figure 4 shows the dendrograms for the two periods. One notable difference between the two dendrograms is that e-business and e-commerce, almost interchangeable in the first period, diverged in the second period.

Lastly, this approach is scalable. The study has examined the six magazines for 50 ITs over ten years, already surpassing the scale and scope of many IT classification studies. While we have used six magazines for this illustration, automation in this approach is not limited in the type or number of discourse outlets or the type or number of technologies.

5.3 Limitations and Future Research

The benefits we just discussed can be realized only within the limitations of this approach. First, the 50-IT list, despite the diversity in it, is an *ad hoc* list that we generated based on our own knowledge of the various domains in IT. Future research should develop a more systematic way to identify technology categories to be included in a taxonomy. While it is never our intention to exclude human knowledge from the selection process, we suggest using automated topic detection techniques to generate a preliminary list and then developing criteria for selection by humans. Second, the quality of the taxonomy must be assessed against "ground truth," which is currently absent in our approach. Therefore, a logical next step is to search for or develop baselines for evaluating quality. Lastly, the usefulness of a taxonomy will ultimately be determined by how well it satisfies users' requirements, which vary across user groups such as IT managers and IT researchers. Consequently, future research should collect requirements from target user groups, build taxonomies according to specific requirements, and test usability in different user groups.

6. CONCLUSION

In conclusion, our combined use of co-occurrence analysis, hierarchical clustering, and multidimensional scaling has given rise to a representative, dynamic, and scalable approach to building IT taxonomies. Properly developed taxonomies are useful in many aspects of IT management. For providers of IT products and services, a taxonomy empirically developed may complement the product categories designated in a top-down design process. For adopters of IT products and services, taxonomies are needed for portfolio management. For scholars in the iField, our approach not only helps many make sense of the complex and dynamic relationships among numerous technologies, but also bridges two related, but thus far largely separate streams of research in iSchools: information management and IT management. As we have shown, commonplace information management techniques such as co-occurrence analysis and clustering can be profitably integrated and applied to solve problems in IT management. Hence the moral of this study: There is a large amount of information about a large amount of IT. A large amount of IT generates a large amount of information. Therefore, effective IT management and effective information management take place hand in hand.

7. ACKNOWLEDGMENTS

This paper is based upon work supported by the National Science Foundation under Grants No. IIS-0729459 and SBE-0915645.

8. REFERENCES

[1] Barki, H. and Rivard, S. 1988. An information systems keyword classification scheme. MIS Quarterly. 12, 2, 299-322.

[2] Barki, H., Rivard, S., and Talbot, J. 1993. A keyword classification scheme for is research literature: An update. MIS Quarterly. 17, 2, 209-226.

[3] Burgess, C. and Lund, K. 1997. Parsing constraints and high-dimensional semantic space. Language and Cognitive Processes. 12, 177-210.

[4] Burgess, C. and Lund, K. 1997. Representing abstract words and emotional connotation in high-dimensional memory space. In Proceedings of the 19th Annual Conference of the Cognitive Science Society (Mahwah, NJ). Lawrence Erlbaum Associates. Inc., 61-66.

[5] Chandon, J.-L. and Pinson, S. 1981 Analyse typologique: Théories et applications. Masson.

[6] Dwivedi, Y., Mustafee, N., Williams, M. D., and Lal, B. 2009. Classification of information systems research revisited: A keyword analysis approach. In Proceedings of Pacific Asia Conference on Information Systems (Hyderabad, India).

[7] Ein-Dor, P. and Segev, E. 1993. A classification of information systems: Analysis and interpretation. Information Systems Research. 4, 2, 166-204.

[8] Everitt, B., Landau, S., and Leese, M. 2001 Cluster analysis. Oxford University Press.

[9] Farbey, B., Land, F. F., and Targett, D. 1995. A taxonomy of information systems applications: The benefits' evaluation ladder. European Journal of Information Systems. 4, 1, 41-50.

[10] Fichman, R. G. 2004. Going beyond the dominant paradigm for information technology innovation research: Emerging concepts and methods. Journal of the Association for Information Systems. 5, 8, 314-355.

[11] Fiedler, K. D., Grover, V., and Teng, J. T. C. 1996. An empirically derived taxonomy of information technology structure and its relationship to organizational structure. Journal of Management Information Systems. 13, 1, 9-34.

[12] Hansen, P. and Jaumard, B. 1997. Cluster analysis and mathematical programming. Mathematical programming. 79, 1-3, 191-215.

[13] Harris, K., Caldwell, F., Linden, A., Knox, R., and Logan, D., "Taxonomy creation: Bringing order to complexity," Gartner, Inc. QA-20-8719, 10 September 2003.

[14] Jagerman, E. J. 2006 Creating, maintaining and applying quality taxonomies. Lulu.com.

[15] Kruskal, J. B. 1977. The relationship between multidimensional scaling and clustering. In Classification and clustering, J. Van Ryzin, Ed. Academic Press, New York, 17-44.

[16] Lambe, P. 2007 Organising knowledge: Taxonomies, knowledge and organisational effectiveness. Oxford: Chandos.

[17] Logan, D., "Best practices for taxonomy creation," Gartner, Inc. G00167683, 2009.

[18] Napior, D. 1972. Nonmetric multidimensional techniques for summated ratings. In Multidimensional scaling; theory and applications in the behavioral sciences, R. N. Shepard, et al., Eds. Seminar Press, New York, 157-178.

[19] Schvaneveldt, R. W. 1990 Pathfinder associative networks: Studies in knowledge organizations. Ablex Pub. Corp.

[20] Shepard, R. N. and Arabie, P. 1979. Additive clustering: Representation of similarities as combinations of discrete overlapping properties. Psychological Review. 86, 2, 87-123.

[21] Smadja, F. 1993. Retrieving collocations from text: Xtract. Computational Linguistics. 19, 143-177.

[22] Spence, D. P. and Owens, K. C. 1990. Lexical co-occurrence and association strength. Journal of Psycholinguistic Research. 19, 5, 317-330.

[23] Swanson, E. B. and Ramiller, N. C. 1993. Information systems research thematics: Submissions to a new journal, 1987-1992. Information Systems Research. 4, 4, 299-330.

[24] Takane, Y., Young, F. W., and De Leeuw, J. 1977. Nonmetric individual differences multidimensional scaling: An alternating least squares method with optimal scaling features. Psychometrika. 42, 1, 7-67.

[25] Vessey, I., Ramesh, V., and Glass, R. L. 2005. A unified classification system for research in the computing disciplines. Information and Software Technology. 47, 4, 245-255.

[26] Wang, P. 2009. Popular concepts beyond organizations: Exploring new dimensions of information technology innovations. Journal of the Association for Information Systems. 10, 1, 1-30.

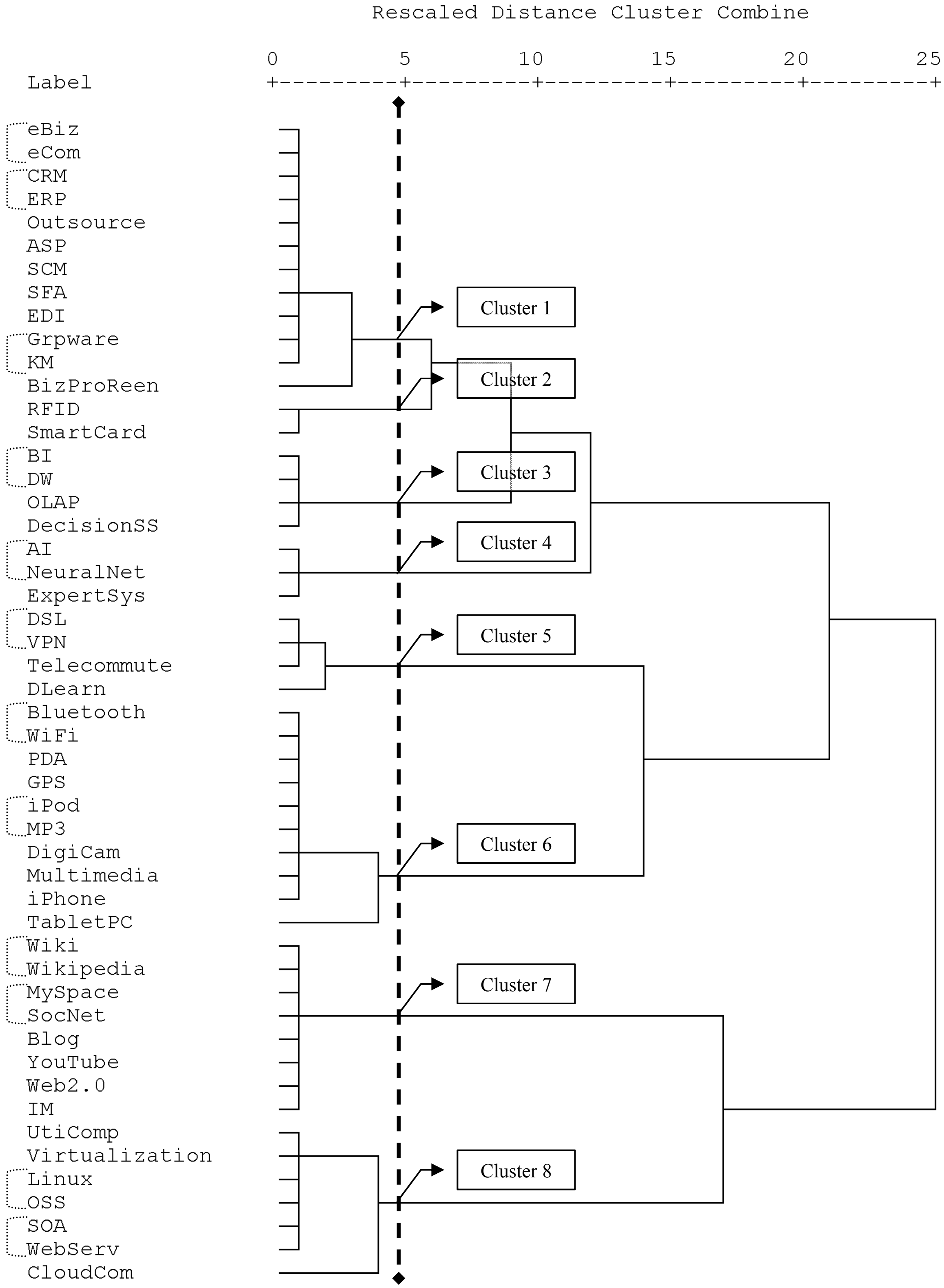


Figure 1. Dendrogram generated from hierarchical clustering analysis

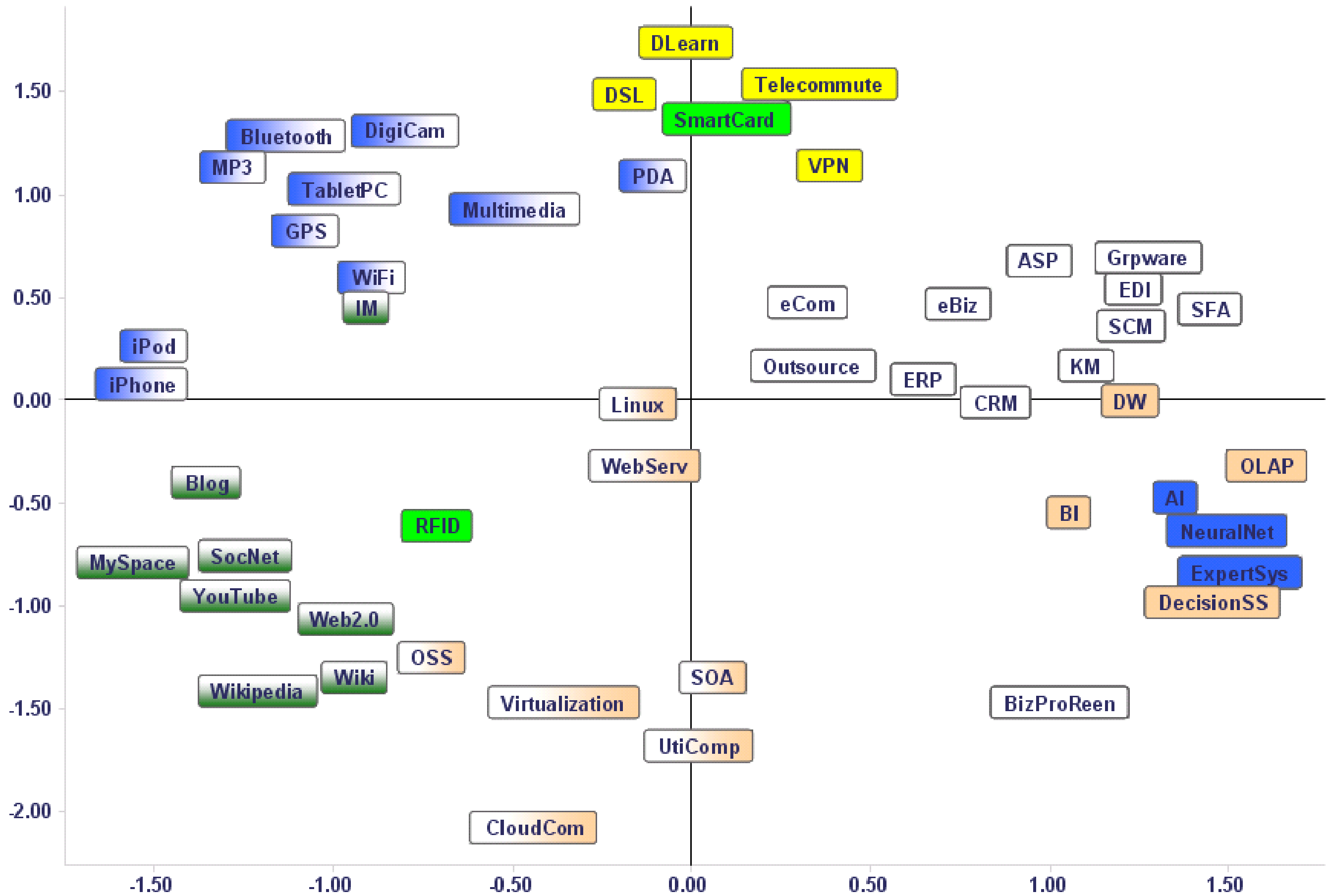
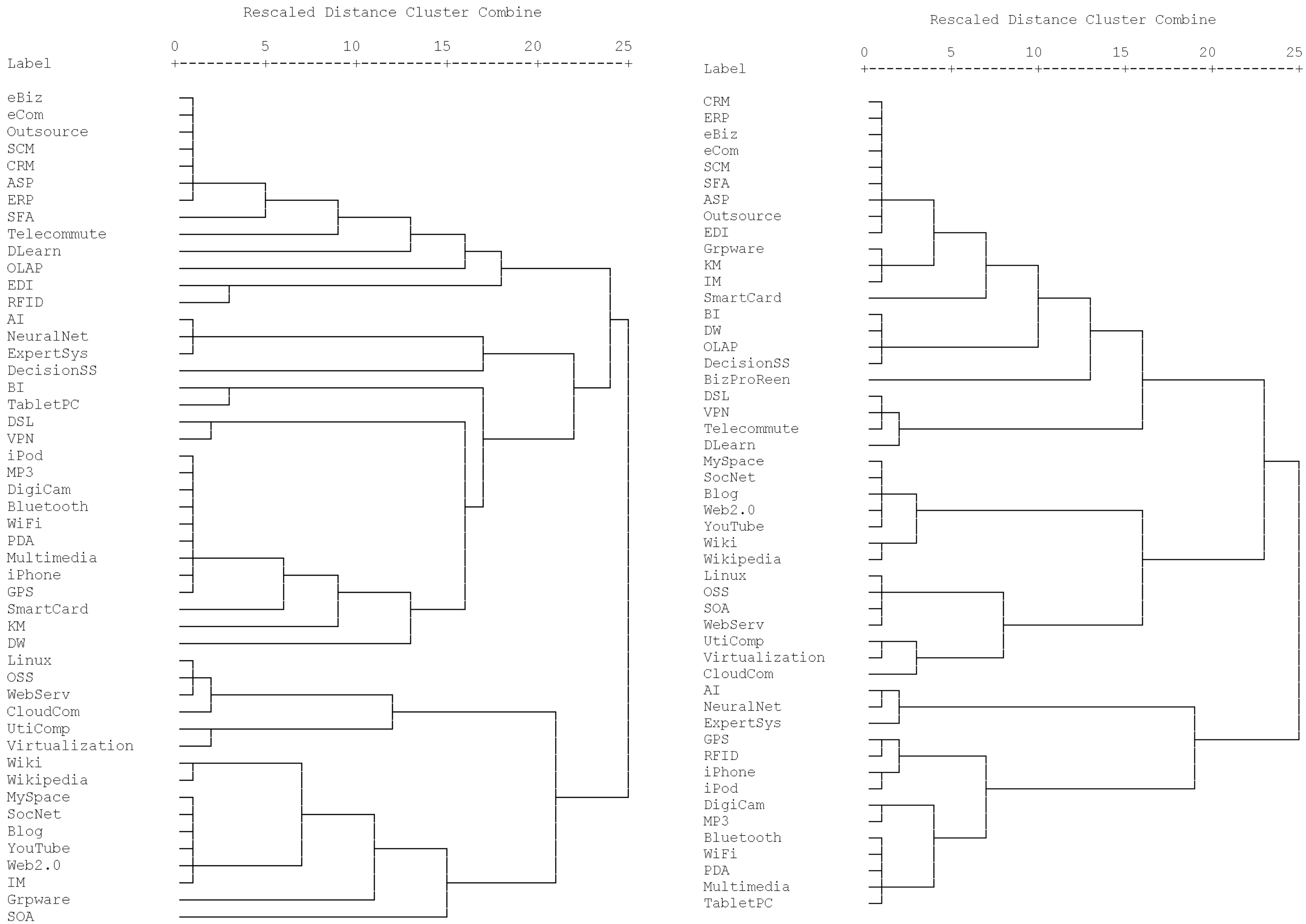
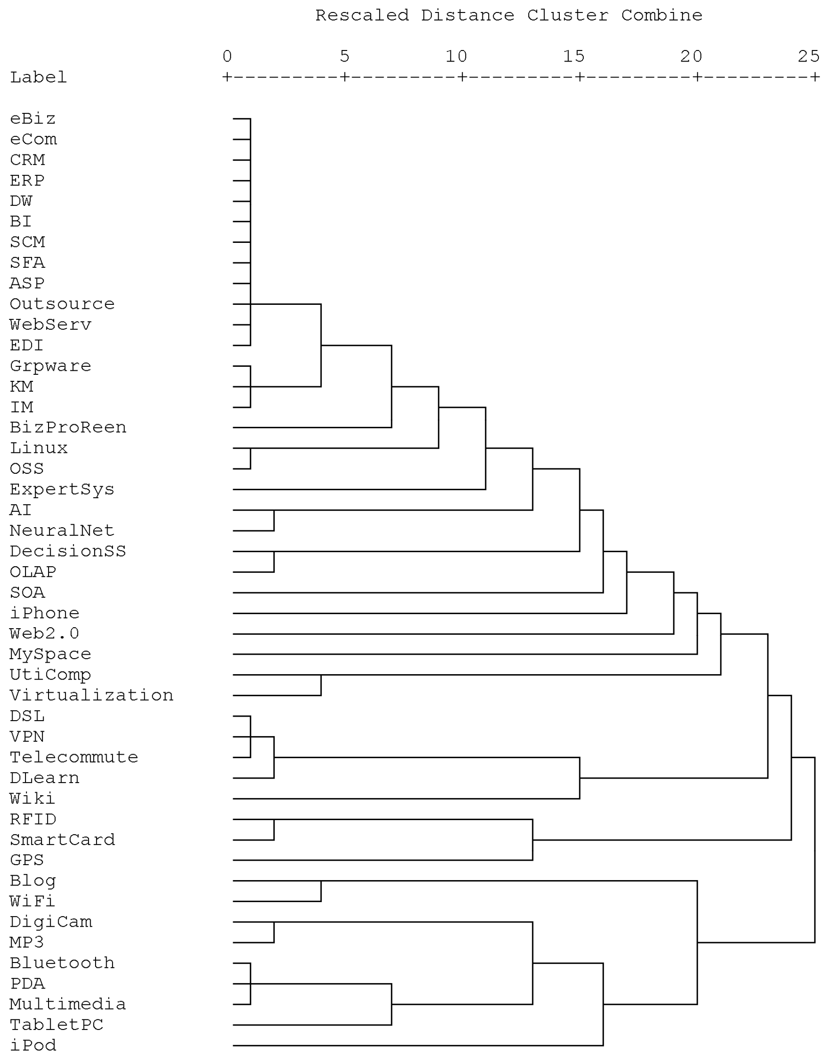


Figure 2. A multidimensional scaling (MDS) plot of the 50 ITs (six magazines, 1998-2007)

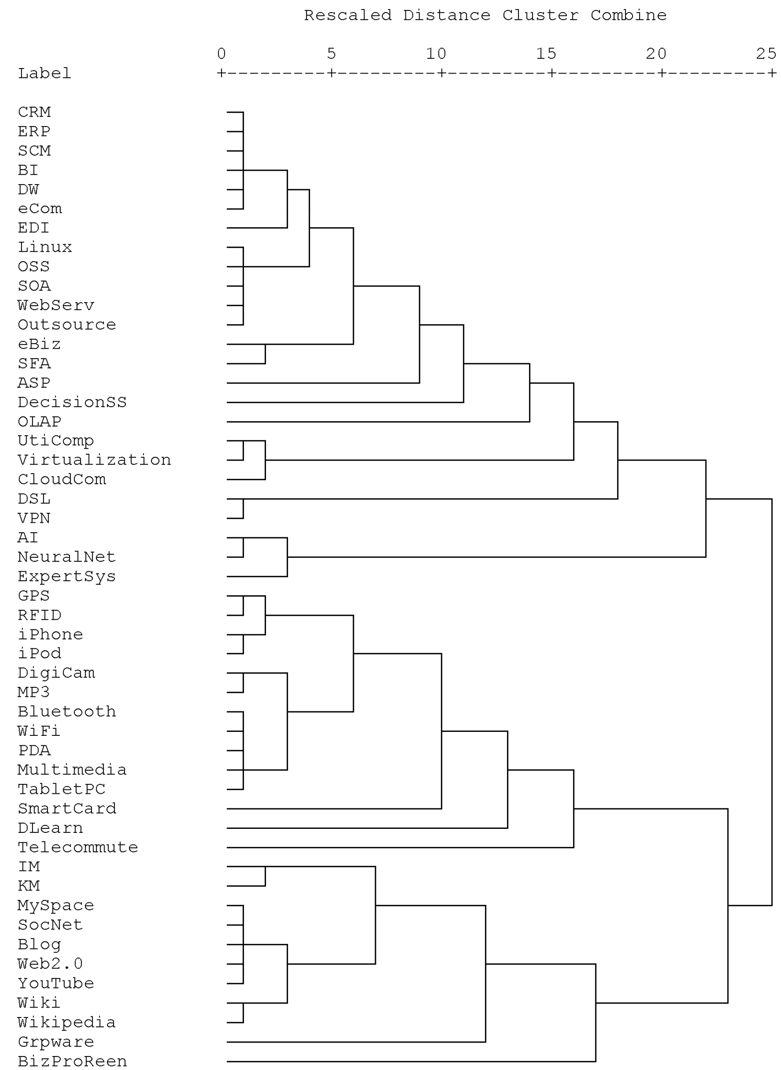


BusinessWeek *InformationWeek*

Figure 3. Dendrograms generated from hierarchical clustering analysis of data in individual magazines (1998-2007)



1998-2002



2003-2007

Figure 4. Dendrograms generated from hierarchical clustering analysis of *InformationWeek* data in 5-year periods