Available online at www.sciencedirect.com
SCIENCE DIRECT®
Technovation 26 (2006) 784–795
technovation
www.elsevier.com/locate/technovation

Linking innovative product development with customer knowledge: a data-mining approach

Chao-Ton Su[a,*], Yung-Hsin Chen[b], D.Y. Sha[b]

[a]Department of Industrial Engineering and Engineering Management, Room 802, Engineering Bldg #1, National Tsing Hua University, Hsin Chu 300, Taiwan, ROC
[b]Department of Industrial Engineering and Management, National Chiao Tung University, Hsin Chu 300, Taiwan, ROC

Abstract

In today's digital economy, knowledge is regarded as an asset, and the implementation of knowledge management supports a company in developing innovative products and making critical management strategic decisions. Product innovation must link technological competence such as engineering and process know-how with knowledge about the customer, so that the product will meet the customers' needs, in order to secure market acceptance. Even though the importance of knowledge management in the technological innovation of a product has long been recognized, its potential for customer knowledge management has not been widely researched.

To address the importance of the need of customer knowledge in innovative product development, this paper proposes an E-CKM model with a methodology for precisely delineating the process of customer knowledge management for innovative product development. In the knowledge management domain, an important task is the conversion of tacit knowledge into explicit knowledge, allowing information technology, such as web-based surveys and data mining to extract customer knowledge from different market segments. An empirical study applying the E-CKM model has been carried out, and it meets the evaluation criteria in a multiple-assessment scheme for showing a satisfactory result. The result is used in the decision making for innovative product development in order to reduce project risk and secure commercial success.

Keywords: Customer knowledge management; Data mining; Innovative product development; Mobile commerce; Web-based market survey

1. Introduction

Technological innovation allows us to cope with increasingly intensive competition when facing challenges from a rapidly changing market situation. Most companies make efforts in knowledge management (KM) to enhance their competitive advantage in product innovation in order to ensure market success. An important component in knowledge management is knowledge creation. Knowledge creation is supported by two key factors: (1) converting tacit knowledge into explicit knowledge, and (2) translating this tacit knowledge of customers or experts into a comprehensible form (Nonaka and Konno, 1998). Elaborating on knowledge work can have innovative outcomes, such as the discovery of new technologies for the development of new products and new processes. For an innovative product to be successful, the product innovation for a company must link technological competence, such as engineering and process know-how, with customer competence such as knowledge of customer needs (Danneels, 2002). The importance of knowledge management in the innovation of product technology has been duly recognized, however, the potential for customer knowledge management has not been studied in any great depth (Grover and Davenport, 2001; Soo et al., 2002), and little discussion has been devoted to the outcomes of knowledge application (Gold et al., 2001; Plessis and Boon, 2004). Thus, among many types of knowledge in a company, product knowledge and customer knowledge fall into the 'crucial' category, because they directly contribute to the competitive advantage and financial performance of the company. Therefore, any study on knowledge work improvement should focus on making products/services more attractive in order to increase value (Davenport et al., 1996).

* Corresponding author. Tel.: +886 3 574 2936; fax: +886 3 572 2932.
E-mail address: ctsu@mx.nthu.edu.tw (C.-T. Su).

0166-4972/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.technovation.2005.05.005

In the digital economy, customer relationship management (CRM) is a contemporary management tool. It manages the relationship with customers by employing up-to-date information technology (IT) such as on-line data analysis, data-mining and database management in order to understand, communicate with, and to attract them. Its objective is to satisfy and retain customers (Dyche, 2002). Increasing the productivity of knowledge work and managing customers' knowledge so as to understand their needs and wants, enables a company to gain a competitive advantage in the market. Recently the 'customer knowledge management (CKM)' model has drawn much attention by the combining of both the technology-driven and data-oriented approaches in CRM and the people-oriented approach in KM, with a view to exploit their synergy potential (Davenport et al., 2001; Garcia-Murillo and Annabi, 2002). The expectation from this endeavor is to more articulately delineate knowledge 'for' customers, knowledge 'about' customers, and knowledge 'from' customers, so that a more beneficial product can be delivered to the right group of customers, to prevent product failure and to ensure commercial success.

With this background and the objective of addressing the essentiality of customer knowledge in innovative new product development (NPD), this paper presents a methodology to support the argument that in order to ensure business excellence, a product's features must meet the needs of specific customer groups in the market. This is accomplished by a target marketing-oriented customer-knowledge management-model implemented by information technology (E-CKM model). With the introduction of a web-based survey approach and a data-mining technique for observing the outcome of customer knowledge, the customer knowledge management process is as follows. First, the features of a product are transformed into 'benefits that customers need', paving the way to understanding the response of the customer toward the benefits those features bring them. Next, a customer's needs toward the perceived product benefits are taken as a basis for forming market segments. In other words, by converting tacit customer knowledge into explicit knowledge, a company can develop different products for various customer groups having a similar attraction.

Market segmentation, market targeting, and market positioning are the three major tasks to be carried out in target marketing (Kotler, 2003). To be successful in market, market segmentation is certainly a very important task. In order to prove that the E-CKM model is applicable in the field, three criteria are employed for its evaluation: (1) Does the customer accept the web-based survey approach and does he render a sufficient response for data mining? (2) Can the data mining techniques successfully extract the customer knowledge in order to facilitate NPD? (3) If multiple data mining techniques are all qualified to cluster customers into segments, then which one is the most appropriate?

The remaining part of this paper is structured as follows. In Section 2 we review the literature on knowledge management regarding innovation, new product development and the customer in the market place. We provide a detailed description of the E-CKM model implementation for managing customer knowledge in innovative product development in Section 3.

In Section 4 we present an empirical study of the E-CKM model methodology for an innovative Telematics product development project. The satisfactory outcome in meeting the evaluation criteria confirms the feasibility of this model in the real world business environment. We have shown how the outcome of customer knowledge management can be applied to make product variants for different market segments, with a view of reducing project risk, meeting customers' satisfaction and improving business success. Finally, we discuss the limitations and contribution of the E-CKM model in Section 5, and we draw our conclusions and indicate directions for further research in Section 6.

2. Knowledge management for innovation, new product development and the customer in the market place

Perter Drucker defines innovation as 'The effort to create purposeful, focused change in an enterprise's economical or social potential' (Drucker, 1998). He indicates that most successful innovation stems from seven areas of opportunities, three of which are: 'industry and market change', 'changes in perception' as well as 'new knowledge'. Knowledge-based innovation requires not merely one kind of knowledge, but many, and the innovation that creates new users and new markets should be carefully aimed at the specific application. Betz (2003) defines technological innovation as: 'Both the invention of a new technology and its introduction into the marketplace as a new hi-tech product, process, or service.' Technological innovation allows us to cope with increasingly intensive competition in a rapidly changing market place. Therefore companies should use their knowledge for improving their competitive advantage in product innovation for market success, by enhancing their capability to manage that knowledge so as to convert it into useful products and services.

Researchers have since defined knowledge category using the concept of explicit knowledge and tacit knowledge (Polanyi, 1966). Nonaka has used a SECI model (SECI: Socialization, Externalization, Combination and Internalization) to depict knowledge creation as a spiral process of interaction between explicit knowledge and tacit knowledge. The 'externalization' step which takes place at 'interaction ba' plays an important role in knowledge creation, and is supported by two key factors: (1) converting tacit knowledge into explicit knowledge and (2) translating the tacit knowledge of customers or experts into comprehensible forms (Nonaka and Konno, 1998). Investment in knowledge work can lead to innovation efforts such as the discovery and the development of new technologies, new products, and new production processes according to Carneiro (2000). Danneels (2002) offers an important insight when he argues that product innovation for a company must link technological competence such as engineering and manufacturing know-how with customer competence such as knowledge of customer needs. Therefore, it is worthwhile to manage the knowledge a company desires to have. It is also imperative that the knowledge management process for transforming knowledge into a company's innovation in products and services is institutionalized as a competitive advantage.

The importance of knowledge for a company's competitive advantage has been well recognized. However, justifying this knowledge as being valuable is a must for a company in order to qualify the knowledge as an intangible asset. 'Knowledge asset underpins competence, and the competence in turn underpins the company's products and services offering in the market.' (Styre, 2002). Indeed, the success of a knowledge-conscious company relies on its efficiency in creating knowledge, and its effectiveness in applying that knowledge to products and services that offer a deliverable value to customers thereby generating a profit for the company. Other researchers have emphasized the use of KM for reducing the risk in NPD, by collecting data from internal and external sources and extracting relevant information in order to prevent product failure. The internal problems affecting product failure include the company's ability to meet product performance, reliability, or cost requirements, while the external problems include: unsuccessful product acceptance in the market, changing regulations, and so on (Cooper, 2003).

Today the role that KM plays in the NPD activities is better understood. However, the role has resulted only in making a contribution to in-company NPD outcomes such as product/service quality, reduced cost, and deliverables-to-market, and not in terms of market outcomes such as sales, customer satisfaction, and return on investment. It is important that both in-company and market outcomes are linked and jointly assessed for real success to be achieved. Therefore, customer knowledge that results in sales is an important attribute of any NPD project.

CRM, via the use of IT, has already been recognized as a contemporary management tool in the digital economy for managing the relationship with customers. It does so by taking advantage of a database management system or on-line data analysis, to assist a company in its management decisions. In order to maintain a good relationship with customers, it is crucial that a company communicates and interacts with its customers in a satisfactory manner, and provides market offerings that continuously meet customers' changing needs. This requires the deliberate management of 'customer knowledge' (Davenport et al., 2001; Garcia-Murillo and Annabi, 2002) such as:

(1) Knowledge 'for' customers
satisfies customers' requirements for knowledge about products, the market, and other relevant items.
(2) Knowledge 'about' customers
captures customers' background, motivation, expectation, and preference for products or services.
(3) Knowledge 'from' customers
understands customers' needs pattern and/or consumption experience of products and/or services.

In this regard, customer knowledge obtained via a CRM system is a valuable intellectual asset for a company to develop or improve products and services in order to meet or even exceed customers' expectations. CRM systems that collect information for customer knowledge are classified into three main categories (Dyche, 2002):

(1) Operational CRM systems
enhances the efficiency of a CRM process through service-center management and marketing-automation like database marketing.
(2) Analytical CRM systems
evaluates knowledge of an individual customer's attitude, needs, and values for cluster analysis. Data-mining is a typical technique in this category.
(3) Collaborative CRM systems
synchronizes customer communication time through channels such as e-mail, the Internet, and/or the telephone.

In the literature most studies on KM and CRM are treated in separate research domains. However, lately their mutual synergy potential has drawn the attention of researchers in the field. By employing KM in an effort to help CRM to transcend from its original technology-driven and data-oriented approach into a more people-oriented 'customer knowledge management' model or CKM model, it has already invoked a convergence of the two (Davenport et al., 2001; Garcia-Murillo and Annabi, 2002). The CKM model emphasizes a bi-directional communication channel. This interaction with customers and customer knowledge management, set up strategies for how a company can develop attractive innovative products, or improve its services to win the satisfaction of its customers.

3. An E-CKM model for customer knowledge management

Customer knowledge management is crucial for NPD project. In order to address this issue, this paper proposes a conceptual framework entitled the E-CKM model by incorporating IT into the CKM model, as shown in Fig. 1.

The proliferation of business applications on the Internet has grown rapidly since the late 1990s. Compared with a conventional mail survey the web serves as an ideal delivery

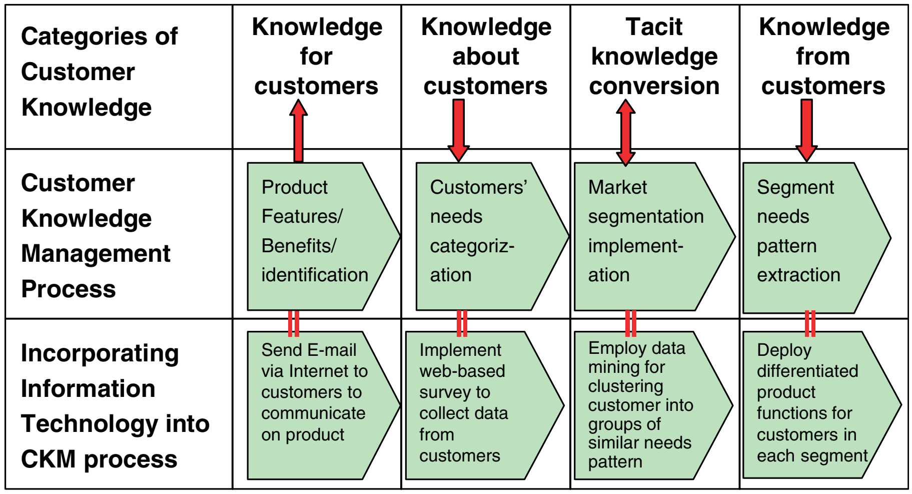

Fig. 1. The E-CKM model as applied in customer knowledge management for innovative product development.

system for a company to conduct an on-line survey because of its timeliness and ease.

However, 'coverage error problem' and 'sample error problem' are the two major factors when considering the validity of this emerging method. Then there is the 'filter problem' which is associated with software tools that filter the so-called 'junk' and 'spam' mails. In addition there is the 'delete problem' for impatient recipients to give up answering the whole survey instrument, and all of this affects the response rate (Fought et al., 2004). Under these circumstances, a company wanting to conduct a web-based survey to extract customers' knowledge, should take measures to address the above-mentioned problems.

Countermeasures include: (1) selecting the target survey samples from company's current customers database, rather than using a general sampling plan, (2) sending a short e-mail to solicit sample recipients to answer a survey instrument posted on a certain website, rather than sending them an e-mail with a long list of questionnaires attached, and so on.

In the E-CKM model, the CKM process comprises four stages which are supported by the applications of different methods in information technology. The implementation scheme is as described below.

3.1. Product features/benefit identification

'Benefit' is defined as the needs a customer wants fulfilled from a product purchase and usage. After a new product concept is formulated, the company identifies perspective product benefits in terms of a customer's perceived value, in the form of features, functions, and other attributes which can be communicated to the customers. At this stage, the company delivers product knowledge 'for' the customers. Individual customers may make response based on their own attitudes toward these features or benefits, via a bi-directional communication channel like personal interviews, phone interviews, self-administered questionnaires, etc. Among them, a mail survey by sending a mail with self-administered questionnaires has been assessed as a good way of accomplishing the objective (Cooper and Schneider, 2001).

In the digital era we live in, information technology can improve communication efficiency and effectiveness by including product knowledge in the self-administered questionnaires in the e-mail to the customers, or by sending out an e-mail to customers soliciting them to answer the survey instrument posted on a predetermined website. The total of the customers' responses, collected in the form of data will go to a server computer, and be arranged as the customer response database. This procedure creates an opportunity to identify what product knowledge the customers have already acquired about the features offered.

3.2. Customers' needs categorization

A market survey instrument such as a questionnaire provides responses from customers about their attitude, preference, needs, and perceived value for product features offered. If a company wishes to offer products to target customers, a procedure has to be developed for grouping customers in the study, by a pattern of needs towards a set of benefits potentially available in the perspective product.

In the E-CKM model, with the aid of IT and using an interval rating scale instrument, an electronic form survey is posted on a website. This is done to collect the primary data for acquiring the personal background of a customer, as well as the extent to which he desires each product feature, in order to extract clues on the needs of the customer. These clues are then categorized into homogeneous groups. In the field of business research, there are several response methods using an interval rating scale. Among them, the 'multiple rating list scale' is the most suitable one for a web-based questionnaire, as it provides a layout of easy visualization (Cooper and Schneider 2001). It also accepts mouse 'click' streams to replace conventional customer responses by hand-written format. In every questionnaire, an individual customer presents his or her own specific pattern of needs, in addition to his or her demographic data and other relevant information. The total of the feedback of the questionnaires collected from all customers in the study first undergoes a data cleaning process, and then becomes the database ready for data mining. This then leads to a categorization of the customers' needs. At this stage, the company acquires knowledge 'about' the customers by understanding the customers' background, needs, and preference pattern toward product features.

3.3. Market segmenting for converting tacit customer knowledge into codified knowledge

Recent studies evaluated the effectiveness of different segmentation bases by considering the following six criteria: identifiability, substantiality, accessibility, stability, actionability and responsiveness. A benefit-based or needs-based segmentation methodology is ranked as the most appropriate if taking into account the overall performance (Kamakura and Wedel, 2000). The fact is that demographic segmentation merely describes customers' behaviour, but fails to explain the reason why it is; while the benefits a product delivers can alter the customers' needs and attitude (Haley, 1985).

Through communication and a web-based survey, a company is able to utilize knowledge 'for' customers and knowledge 'about' customers, and conduct the appropriate market segmentation task. After the segments are formed through data mining, each segment's pattern of needs toward product features is well delineated. Now the different characteristics of each segment can be identified and analyzed. At this stage the tacit customer knowledge, dispersed among the individual customers is excavated, and it can be codified into explicit customer knowledge desired by the company.

3.4. Customers' needs pattern extraction

Once the segmentation task is completed, the characteristics of the customers' needs in each segment are studied in order to extract the needs patterns in each segment. Therefore, the knowledge 'from' customers enables the company to aim at the right target market segments. It also enables them to make the appropriate strategic business decisions in the product variant development plan and marketing activities. It helps the company to revise the original definition of the product, set priorities for product attributes to be developed, enhance the functionality of the attractive product elements, and rule out product features in which the customers show no interest. The different patterns as demonstrated by the different segments also becomes useful information for the company, to make the appropriate decisions for future product line-ups to serve the different target segments. However, the clustering operation in this study is based on the data patterns provided by the customers with the rating scale instrument via a web-based survey questionnaire. For a survey scale instrument to be valid and possess practical business utility, the construct it represents must be sufficiently reliable. In the E-CKM model, a multiple assessment scheme is used to justify both the clustering analysis and the survey construct reliability.

3.5. Data-mining technique to implement market segmentation

This paper proposes three data-mining methods as applied in the E-CKM model: K-means; the unsupervised Self Organizing Feature Map (SOM) neural network (Kohonen, 1990); and a network based on the Fuzzy Adaptive Resonance (FuzzyART) theory (Carpenter et al., 1991). A detailed description of these methods will not be presented in this paper, instead only parameters that affect their operation characteristics and performance are described.

The reason why this study selects K-means, SOM and FuzzyART network as data mining methods is based on the following arguments: (1) K-means algorithm has been regarded as the benchmark tool in evaluating the performance of an artificial neural network. It is the most popular and most widely used procedure for market segmentation (Kamakura and Wedel, 2000), (2) urge on more study by researcher in using unsupervised one like SOM, because cluster-based problems such as market segmentation problems have been successfully tackled by Multi-layered Perceptron (MLP) network, basically a supervised network architecture (Vellido et al., 1999), (3) The very rare application of FuzzyART network in segmentation literatures is found to be due to its complexity, motivating the exploration of its feasibility in the marketing segmentation problem (Chen et al., 2002).

3.5.1. K-means clustering method

K-means is a non-hierarchical clustering method which is normally chosen as the benchmark for other clustering algorithms. The K-means algorithm operation requires choosing a number of clusters and an initial cluster center.

3.5.2. The self organizing map network as one of the clustering methods

The SOM network can learn a topological map from an n-dimensional input into a two-dimensional feature space. It provides a way to visualize high-dimensional data through a two-dimensional presentation (Kohonen, 1990). The network architecture and algorithm have been used to cluster a data set of p continuous-valued vector $\mathbf{X}=(\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots\boldsymbol{x}_i,\ldots\boldsymbol{x}_n)$ into m clusters. In this study, the data set collected from customers' responses are fed into the SOM network to form a given number of clusters. The topological neighborhood parameters and the learning rate must be set for the operation of the SOM network.

3.5.3. The FuzzyART network as one of the clustering methods

There are three parameters to determine the operation of the FuzzyART network: choice parameter, learning rate parameter, and vigilance parameter (Carpenter et al., 1991). For the same input data pattern set with the same value for choice parameter and learning rate parameter, a low vigilance rate will result in a small number of coarse clusters. On the other hand, a high vigilance rate will lead to a large number of finer clusters.

3.5.4. Evaluation criteria in a multiple assessment scheme for a data mining task

In the E-CKM model a multiple assessment scheme has been proposed to deal with the issue of justification. The three evaluation criteria are: (1) Does the customer accept the web-based survey approach and render a response sufficient for data mining? (2) Can data mining techniques successfully extract customer knowledge to facilitate NPD? (3) Which one is the most appropriate technique, if multiple data mining techniques are qualified for clustering customers into segments?

A sufficient survey response sample size of 1 000 is assumed to be the threshold for criteria number 1. Criteria numbers 2 and 3 are concerned with (1) how to justify the K-means, the SOM network, and the FuzzyART network as qualified data mining methods for performing a clustering operation for market segmentation, and (2) in case that all three methods are qualified for the clustering operation, which one will be the most appropriate for the E-CKM model?

The first problem that arises is: how to determine the best number of clusters or the optimal 'natural' number of clusters. Defining a criterion function that measures the clustering quality of any clustering solution, and finding the cluster solution that presents the extreme value of the criterion function is one approach to solve the problem (Duda et al., 2001). In other words, a clustering solution to come up with solution for different numbers of clusters is performed repeatedly, in order to compare the criterion function value for each clustering operation. If there is a 'large gap' in the value of the criterion function among clustering solutions, then this suggests a 'natural' number of clusters. Among many criterion functions for clustering, this study has chosen *R*-squared (RS), the ratio of SS_b (between-clusters variation) to SS_t (total variation), as the criterion function, and $SS_t = SS_b + SS_w$, where SS_w is the within-cluster variation. For a given data set, the greater the between-cluster variation, the more internally homogeneous each cluster will be. Plotting the *R*-squared values as a vertical axis against the number of clusters as a horizontal axis, an 'elbow' point in the plot indicates the best clustering solution or the optimal number of clusters (Sharma, 1996).

The second problem will be resolved by adopting the existing market research paradigm: using the most commonly-used reliability coefficient for a rating scale survey instrument, namely, Cronbach's alpha coefficient (Cronbach, 1951; Peterson, 1994), an estimator of internal consistency for a survey instrument, which is approximately defined by the formula:

$$\alpha = \left(\frac{K}{K-1}\right)\left(1 - \sum_{i=1}^{k} \partial_i^2 / \partial_s^2\right) \tag{3.1}$$

where K is the number of items in the survey instrument scale, ∂_i^2 is the variance of item i, and ∂_s^2 is the variance of the whole scale in the survey instrument. The recommendation for the minimally acceptable reliability level expressed by Cronbach's alpha is proposed as: below 0.6 is the unacceptable level; 0.7 is the low level; 0.8–0.9 is the moderate to high level, and 0.9 is the high level (Murphy and Davidshofer, 1988).

4. Case study

4.1. Telematics is an innovative product for mobile commerce

In this era of the digital economy, companies have benefited from the capabilities of the Internet to operate sales and communication channels reaching customers in large geographical regions. This new approach is called the 'electronic-commerce (e-commerce)'-selling and buying of products/services over the web. The connectivity in e-commerce in terms of fixed user location within the wired communication infrastructure can be further enhanced by innovative 'ubiquitous networking' wireless communication technology to construct 'mobile-business (m-business) or mobile-commerce' under the definition of 'm-business = Internet + wireless + e-business' as given by Kalakota and Robinson (2002).

To demonstrate how to meet future demands of mobile-commerce, an innovative NPD project is initiated in Taiwan. The objective of this project is to explore the opportunity for Taiwan's hi-tech industry, especially the computer and wireless communication hardware manufacturers to penetrate into a new market, specifically the automobile Telematics computer market. The purpose of this NPD project is to serve prospective users that start from Taiwan and then extend to a global market coverage by taking advantage of the economies of scale and economies of scope. The web-based market survey supporting the E-CKM model has been applied in this innovative NPD project to demonstrate that technological innovation requires not only the complexity of knowledge, but also the integration of customer knowledge.

Telematics is the combination of '*Tele*communication' and 'infor*matics*', and is regarded as an innovative technology combining wireless communication, in-vehicle information system, and in-vehicle multimedia computing system. A Telematics-enabled vehicle offers customers a variety of benefits or values such as: enhanced safety and security, as well as convenience, travel information services, entertainment, etc. This allows the vehicle producer to transform from a product manufacturer into an integrated service provider (Funk, 2001). The technology of Telematics augments the conventional automotive functions and features to an elevated level by providing the 'ubiquitous networking', enabling connectivity anywhere, any time, and anyhow. Telematics computer can be considered as an innovative product, when taking into account the benefits rendered, the sophistication of the engineering efforts, the complexity of its application, and the impact on society.

4.2. Data collection

The company prepared a roster of customers randomly selected from an automobile manufacturer's current customer database. It then sent an e-mail to solicit these selected customers to answer a survey instrument posted on a website. After data cleaning, 1 742 effective questionnaires became available. All data were processed by a MySQL database management system installed in an Apache web server running under a Linux operating system environment.

4.3. Questionnaire used in the survey instrument

On the front page of the website appears a brief letter to thank the customers for taking part in the survey, followed by a detailed instruction on how to answer the questions. The first part of the questionnaire is to determine the demographics of the customers, including age (younger than 20 years old, between 20 and 30, between 30 and 40, between 40 and 50, above 50), gender (male or female), profession (12 kinds), educational background (below high school, high school, college, university, graduate school), income (7 ranges), marital status, and number of children. To make it easy to answer the questions, the numerical data was arranged in increasing order of certain predetermined ranges, for example:

Age: • under 20, • 20 to 30, •30 to 40, • 40 to 50, • over 50

The small circle in front of each range is for customers to 'click' with the mouse and indicate their reply to a particular question.

The second part of the questionnaire in the main survey instrument lists 29 items of Telematics product feature that delivers certain benefits to customers. The 29 items are arranged into 8 groups, and each item is presented in the format shown in Table A1 in Appendix A.

The third part collects information of a customer's perceived value regarding certain Telematics features. Some major features are selected from the above 29 items. They are presented with five options each, and each with a different monetary value. The customer then clicks the small circle in front of a particular item to disclose his/her perceived value, or the monetary value he/she is willing to pay for one option. This completes the selection task.

4.4. Data processing and analysis

The data set in the database consists of 1472 records of customer's extent of needs which composed of 29 attributes data items. The data item in each record relating to each attribute is one of continuous-valued variables in the set of (0.1, 0.3, 0.5, 0.7, 0.9).

The SAS (Statistical Analysis System) software package is used to carry out the K-means clustering operation. The NeurolShell2 software package is employed to perform clustering in the SOM neural network. Clustering in the FuzzyART network is by means of ART GALLERY, a software package developed by Dr Lars Hasso Linden of Boston University, US.

4.5. Results

4.5.1. Determining the optimal clustering solution-the number of 'natural' clusters

From a pragmatic standpoint of need-based or benefit-based segmentation task, a selection from between three to seven as the number of clusters makes marketing people consider it to be manageable in business practices (Haley 1985). Therefore, the data set is first clustered into three, four, five, six, and seven segments respectively by K-means, SOM, and FuzzyART. Then the *R*-squared value corresponding to each clustering solution can be computed. Tables B1–B3 in Appendix B list the parameters and resultant *R*-squared value for different data-mining methods.

When plotting *R*-squared values as the vertical axis against the number of clusters as the horizontal axis, an 'elbow' point in the plot indicates the best clustering solution or the optimal number of clusters (Sharma, 1996). The 'larger gap' in the value of *R*-squared happens where the number of clusters is five for all three data-mining methods, which suggests that five is a 'natural' number of clusters or market segments. Comparing the *R*-squared values and plots of these three methods, it is evident that the performance of the SOM network is at the same level as that of K-means, while FuzzyART distinguishes itself among these clustering methods in terms of ease of identification.

4.5.2. Determining the most appropriate clustering method

The FuzzyART network, SOM network, and K-means clustering methods result in the same 'natural' number of clusters of five segments as shown in Figs. D1–D3 in Appendix D. However, the issue that remains is: how reliable are the results of these clustering methods? The clustering operation in this paper is based on the needs pattern database as per the response from customers via a web-based survey questionnaire. And, because a survey construct representing a particular segment should be valid and possess practical utility, it must be justified as being reliable by meeting a minimally acceptable level expressed as Cronbach's alpha coefficient.

The Cronbach's alpha coefficient displayed at Table C1 in Appendix C indicates that only clustering solutions by the FuzzyART network fulfill the minimal reliability acceptance level for customers' needs patterns assigned in the rating scale instrument. Although the SOM network and K-means clustering algorithm are both qualified as an adequate method to find out the 'natural' clusters, they unfortunately fail in providing a moderately reliable construct of each segment for meeting the business practice requirement.

Using the multiple assessment scheme, the company selected the clustering result by FuzzyART network to further investigate the needs pattern of customers in each of the five segments. The different needs patterns demonstrated by the different segments therefore become useful information for the company to make the appropriate decisions, for product variant development as well as for a future product line-ups to serve each of the target segments.

5. Discussion

This paper begins with the argument that: (1) product innovation must link with the knowledge of customers' needs, (2) the results of knowledge application should be discussed more fully, (3) information technology such as CRM, web-based survey approaches and data-mining techniques may facilitate the knowledge process. In order to address the importance of customer knowledge in innovative product development, we proposed a target marketing-oriented E-CKM model, by presenting the taxonomy of customer knowledge, namely, knowledge 'for' the customer, knowledge 'about' the customer and knowledge 'from' the customer, as well as the process of how to create customer knowledge in each stage.

A point of interest is the application of the IT-based procedures, such as web-based survey and data mining for a real world innovative NPD project, in the attempt to diminish the 'knowing-doing gap' that has been criticized by major KM research articles. In addition three evaluation criteria are also proposed to scrutinize the robustness of the proposed E-CKM model.

The empirical data collected from a web-based market survey provides considerable support for the E-CKM model, taking into account the evaluation result based on some of the criteria. It is evident from the results that the methodology proposed in this paper can be successfully applied in a business practice by linking NPD with customer knowledge. The approach discussed in this paper clearly delineate the process and procedures of the E-CKM model and puts the research results of the academic community onto a concrete methodology for business practice, allowing for a positive impact in the foreseeable future.

Although the outcome of knowledge application in this study is satisfactory, some inherent limitations must be noted. First, it focused on a product innovation in the automotive industry. The benefits offered by the innovative product concerned only car users, rather that the general public. Second, the findings in this study relied on a quantitative assessment via the needs pattern of an aggregate customers' response. Therefore, it must be assumed that the feedback of sample customers, does represent the needs pattern of all the customers registered in the customer database. Third, the outcome of the customer knowledge application should be supported by a successful clustering method. Instead of using a conventional measurement criteria to determine the most appropriate clustering result, this study employed multiple assessments such as R-squared and Cronbach's alpha coefficient for making a decision, so as to better justify the validity of the application result. This new approach should be subjected to further investigation, even though it provides a pragmatic solution when using a multiple item rating scale to build the survey instrument.

Consequently, the above limitations constrain the generality of the E-CKM model for use in other industries and product contexts. Nevertheless, this study makes a contribution towards the creation of a new methodology by linking innovative product development with customer knowledge, in order to reduce project risk and ensure market success.

6. Conclusion

This article proposed the incorporation of a data-mining technique into the CKM model. The E-CKM model has broken ground for a new methodology in the field of research on customer knowledge management. It does so by using a well-structured process and procedure, aimed at making a contribution in both the academic community and the real world business environment. However, the approach of hybridization of methodologies from different disciplines is still in the incubation stage, requiring a sustainable input from researchers of related fields to collaboratively explore more possibilities in the future. Further research is recommended in several directions:

(1) The hybridization of an unsupervised neural network with a supervised one could use the clustering solution done by the FuzzyART network as the target

for supervised networks, such as the MLP network or the FuzzyART-MAP network.

(2) The application of the E-CKM model in the domain of other components of company knowledge (such as industry knowledge, operations knowledge, supplier knowledge, and competitor knowledge), to explore the influence of other types of knowledge on innovative product development, in order to gain a more competitive advantage.

Appendix A

See Table A1.

Appendix B

See Tables B1–B3.

Appendix C

See Table C1.

Appendix D

See Figs. D1–D3.

Table A1
The main part of the questionnaire: multiple rating list scale for web-based survey on customer needs about Telematics features

Feature item\the extent of needs	Not much needed	Moderately needed	Very much needed	Strongly needed	Extremely needed
Group 1: Automatic route guidance functions					
1. GPS navigation with electronic map	•1	•2	•3	•4	•5
2. Shortest path search	•1	•2	•3	•4	•5
3. Turn and branch prompt	•1	•2	•3	•4	•5
4. Gas station/parking lot position	•1	•2	•3	•4	•5
Group 2: Traffic information					
5. Periodic radio broadcasting	•1	•2	•3	•4	•5
6. Real time information access	•1	•2	•3	•4	•5
Group 3: Emergency services					
7. SOS message in emergency	•1	•2	•3	•4	•5
8. Vehicle tow notification	•1	•2	•3	•4	•5
9. Stolen vehicle tracking	•1	•2	•3	•4	•5
10. Roadside assistance services	•1	•2	•3	•4	•5
Group 4: Travel information					
11. Tourist information guide	•1	•2	•3	•4	•5
12. Travel route information	•1	•2	•3	•4	•5
13. Flight, train, bus schedule	•1	•2	•3	•4	•5
Group 5: Lifestyle and information access					
14. Shopping, on-sales information	•1	•2	•3	•4	•5
News, stock, sports, weather	•1	•2	•3	•4	•5
16. Concierge service	•1	•2	•3	•4	•5

Table A1 (*continued*)

Feature item\the extent of needs	Not much needed	Moderately needed	Very much needed	Strongly needed	Extremely needed
17. Calendar, organizer, address note	•1	•2	•3	•4	•5
18. Voice recording	•1	•2	•3	•4	•5
19. Data synchronization with PDA or notebook computer	•1	•2	•3	•4	•5
Group 6: Mobile commerce					
20. In-vehicle ticket reservation	•1	•2	•3	•4	•5
21. In-vehicle on-line shopping	•1	•2	•3	•4	•5
22. E-mail and short message transceiving	•1	•2	•3	•4	•5
23. Internet web browsing	•1	•2	•3	•4	•5
Group 7: In-vehicle entertainment					
24. DVD, CD, MP3, TV enjoyment	•1	•2	•3	•4	•5
25. Electronic game playing	•1	•2	•3	•4	•5
26. Karaoke singing device	•1	•2	•3	•4	•5
Group 8: Human-machine interface					
27. By voice command	•1	•2	•3	•4	•5
28. By touch screen	•1	•2	•3	•4	•5
29. By joystick	•1	•2	•3	•4	•5

Table B1
Parameters and *R*-squared value for different clustering solutions using the FuzzyART network

No. of segments	No. of epochs	Choice parameter	Learning rate parameter	Vigilance parameter	*R*-Squared
3	1	0.1	0.98	0.14	0.1406
4	1	0.1	0.99	0.16	0.1221
5	1	0.1	0.99	0.18	0.2269
6	1	0.1	0.99	0.20	0.1797
7	1	0.1	0.98	0.21	0.1637

Table B2
Parameters and *R*-squared value for different clustering solutions using the self organizing map network

No. Of segments	No. of epochs	Initial no. of neighborhoods	Initial weight	Initial learning rate	Final learning rate	*R*-Squared
3	5000	2	0.3	0.5	0.0001	0.3882
4	5000	3	0.3	0.5	0.0001	0.4219
5	5000	4	0.3	0.5	0.0001	0.4558
6	5000	5	0.3	0.5	0.0001	0.4572
7	5000	6	0.3	0.5	0.0001	0.4886

Table B3
R-Squared value for different clustering solutions using the K-means algorithm

No. of segments	*R*-Squared
3	0.3877
4	0.4221
5	0.4552
6	0.4662
7	0.4810

Table C1
Cronbach's alpha coefficient for clustering solutions by FuzzyART network, SOM network, and K-means

Cronbach's Alpha for different clustering solutions								
		1st Segment	2nd Segment	3rd Segment	4th Segment	5th Segment	6th Segment	7th Segment
Clustering into 3 segments	FuzzyART	0.9270	0.9389	0.9098				
	SOM	0.7947	0.5864	0.8408				
	K-means	0.5812	0.7921	0.8454				
Clustering into 4 segments	FuzzyART	0.9475	0.9600	0.9640	0.9319			
	SOM	0.7582	0.3148	0.4193	0.8030			
	K-means	0.6921	0.6630	0.8075	0.8494			
Clustering into 5 segments	FuzzyART	0.8531	0.8031	0.9142	0.9287	0.9280		
	SOM	0.7786	0.5752	0.6387	0.4439	0.7954		
	K-means	0.4673	0.5745	0.6430	0.8013	0.7835		
Clustering into 6 segments	FuzzyART	0.9178	0.8932	0.8953	0.8816	0.8825	0.9400	
	SOM	0.7910	0.6230	0.5624	0.4354	0.5141	0.7145	
	K-means	0.6128	0.5043	0.7231	0.7940	0.7819	0.5762	
Clustering into 7 segments	FuzzyART	0.9314	0.9013	0.9371	0.8934	0.8627	0.9359	0.9313
	SOM	0.8027	0.6134	0.4992	0.4987	0.4902	0.4124	0.7189
	K-means	0.5666	0.5149	0.5539	0.8182	0.6810	0.7552	0.6026

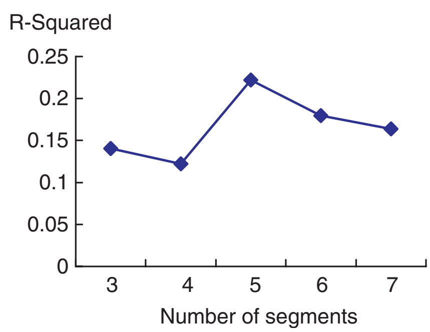

Fig. D1. Clustering solutions for FuzzyART.

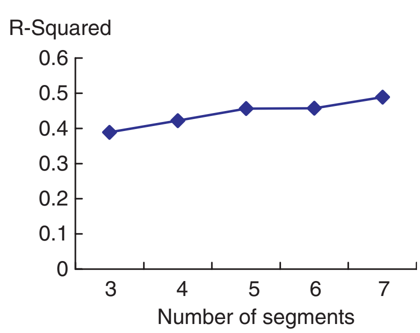

Fig. D2. Clustering solutions for SOM.

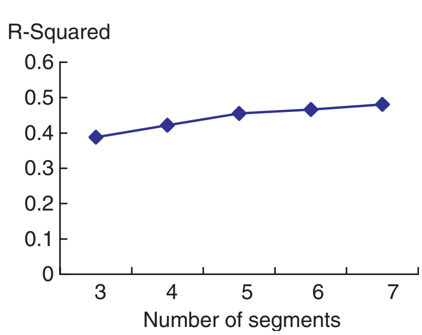

Fig. D3. Clustering solutions for K-means.

References

Betz, F., 2003. Managing Technological Innovation, second ed. Wiley, Hoboken, NJ pp. 22.

Careiro, A., 2000. How does knowledge management influence innovation and competitiveness? Journal of Knowledge Management 4 (2), 87–98.

Carpenter, G.A., Grossberg, S., Rosen, D.B., 1991. FuzzyART: Fast stable learning and categorization of analog pattern by an adaptive resonance system. Neural Network 4, 759–771.

Chen, C.H., Khoo, L.P., Yan, W., 2002. A strategy for acquiring customer requirement patterns using laddering technique and ART2 neural network. Advanced Engineering Informatics 16, 29–240.

Cooper, L.P., 2003. A research agenda to reduce risk in new product development through knowledge management: a practitioners perspective. Journal of Engineering and Technological Management 20, 117–140.

Cooper, D.R., Schindler, P.S., 2001. Business Research Methods, seventh ed. Irwin-Hill Inc, New York pp. 312–318.

Cronbach, L.J., 1951. Coefficient alpha and the internal structure of test. Phychometrika 16 (3), 297–334.

Danneels, E., 2002. The dynamics of product innovation and firm competences. Strategic Management Journal 23, 1095–1121.

Davenport, T.H., Jarvenpaa, S.L., Beers, M.C., 1996. Improving knowledge work process. Sloan Management Review Summer, 53–65.

Davenport, T.H., Harris, J.G., Kohli, A.K., 2001. How do they know their customers so well? MIT Sloan Management Review Winter, 63–73.

Drucker, P.F., 1998. The discipline of innovation. Harvard Business Review, 149–157.

Duda, K.O., Hart, P.E., Stork, D.G., 2001. Pattern Classification, second ed. Wiley, New York pp. 542.

Dyche, J., 2002. Handbook: A Business Guide to Customer Relationship Management. Addison-Wesley Inc, Massachusetts pp. 3–18.

Fought, K.S., Whitten, D., Green, K.W., 2004. Doing survey research on the internet: yes, timing does matter. Journal of Computer Information Systems (XLIV 3), 26–34.

Funk, J., 2001. The Future of the Automotive Industry: Challenge and Concept for the 21st Century. Society of Automotive Engineers, Inc, Warrendale, PA pp. 45–152.

Garcia-Murillo, M., Annabi, H., 2002. Customer knowledge management. Journal of Operational Research Society 53, 875–884.

Gold, A.H., Malhotra, A., Segars, A.H., 2001. Knowledge management: an organizational capabilities perspective. Journal of Management Information System 18 (1), 185–214.

Grover, V., Davenport, T.H., 2001. General perspectives on knowledge management: fostering a research agenda. Journal of Management Information System 18 (1), 5–21.

Haley, R.I., 1985. Developing Effective Communication Strategy: A Benefit Segmentation Approach. Wiley, New York pp. 125–127, 224.

Kalakota, R., Robinson, M., 2002. M-business: The Race to Mobility. McGraw-Hill, New York pp. 6–10.

Kamakura, W., Wedel, M., 2000. Segmentation Bases: Marketing Segmentation—Conceptual and Methodological Foundation, second ed. Kluwer, Massachusetts pp. 7–16.

Kohonen, T., 1990. The self-organizing map. Proceedings of IEEE 78 (9), 1464–1480.

Kolter, P., 2003. Marketing Management, 11th ed. Pearson Education Inc, Uppersaddle River pp. 278–286.

Murphy, K.R., Davidshofer, C.O., 1988. Phychological Testing: Principle and Applications. Prentice-Hall Inc, Englewood Cliffs pp. 89.

Nonaka, I., Konno, N., 1998. The concept of 'ba': building a foundation for knowledge creation. California Management Review 40 (3), 40–54.

Peterson, R.A., 1994. A meta-analysis of cronbach's coefficient alpha. Journal of Consumer Research 21, 387–391.

Plessis, M., Boon, J.A., 2004. Knowledge management in e-business and customer relationship management: South African case study findings. International Journal of Information Management 24, 73–86.

Polanyi, M., 1966. The Tacit Dimension. Routledge and Kegan Paul, London.

Sharma, S., 1996. Applied Multivariate Techniques. Wiley, New York pp. 194–217.

Soo, C., Devinney, T., Midgley, D., Deering, A., 2002. Knowledge management: philosophy, process, and pitfalls, california. Management Review 44 (4), 129–150.

Styre, A., 2002. The knowledge-intensive company and the economy of sharing: rethinking utility and knowledge management. Knowledge and Process Management 9 (4), 228–236.

Vellido, A., Lisboa, P.J.G., Vaughan, J., 1992. Neural network in business: a survey of application (1992–1998). Expert Systems with Applications17, 51–70.

Chao-Ton Su is Professor of Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Taiwan. He received his PhD in Industrial Engineering from University of Missouri—Columbia, USA. From 1985 to 2004, he was a faculty member at the Department of Industrial Engineering and Management, National Chiao Tung University, Taiwan. His research activities include quality engineering and management, operation management and neural networks in industrial applications.Dr Su currently serves as the Editor-in-Chief for *Journal of Quality* of the Chinese Society for Quality (CSQ). He is a member of IIE, CIIE and CSQ and a senior member of ASQ. Dr Su has received two-time (2000–2001 and 2002–2005) Outstanding Research Award from the National Science Council, Taiwan, ROC. In 2001, Dr Su also obtained the Individual Award of the National Quality Awards of the Republic of China.

Yung-Hsin Chen is a PhD Candidate at Department of Industrial Engineering and Management, National Chiao Tung Univsersity, Taiwan. He received BS in Mechanical Engineering and MBA at National Cheng Kung University, Taiwan and MS in Industrial Engineering at Purdue University, West Lafayette, Indiana, USA. Researchwise, he is involved with electrical, electronic and computer technologies applications in automotive industry.Mr Chen currently serves as the General Manager, Division of Electrical and Electronic, at China Motors Corporation, Taiwan. He is a member of Society of Automotive Engineers (SAE). Mr Chen received the Award of 'Young Distinguished Engineer' metal from the China Society of Mechanical Engineering, Taiwan, Republic of China.

D.Y. Sha received his PhD in Industrial Engineering from the University of Taxes at Arlington, USA. He is a professor at the Department of Industrial Engineering and Management, National Chiao Tung University, Taiwan ROC. His research interests include production/operations management and systematic innovation. Professor she was awarded the Industrial Engineer metal from the Chinese Institute of Industrial Engineers, Taiwan, ROC. He is a senior member of IIE and a litetime member of CIIE and CSQ.