Incorporating the Knowledge Management Cycle in E-Business

Ying-Lie O
Subdepartment of Information Systems, Department of Computer Science
University of Twente, The Netherlands
Email: ying@cs.utwente.nl

Abstract

In e-business, knowledge can be extracted from the recorded information by intelligent data analysis and then utilised in the business transaction. E-knowledge is a foundation for e-business. E-business can be supported by an intelligent information system that provides intelligent business process support and advanced support of the e-knowledge management cycle. Knowledge is stored as knowledge models that can be updated in the e-knowledge management cycle. As illustrated in examples, the e-knowledge cycle aids in the business decision taking, production management, and costs management.

Keywords

e-business, intelligent data analysis, intelligent information systems, knowledge management.

INTRODUCTION

Knowledge management is an issue for both enterprise organisation and management, as well as the support of advanced information technology. More specifically, enterprise knowledge management [10] entails formally managing knowledge resources in order to facilitate access and reuse of knowledge. Knowledge management is becoming a critical success factor for enterprises [11].

A business model generally describes the way business is performed according to the business goals and strategies. The business process model consists of activities that support the business model. A *business transaction* is a chain of business processes associated with business value, exchange of information between customer, business, and other involved parties, and product handling from the beginning to the end. Knowledge management can be associated to the business transaction by indicating the use and creation of knowledge in the business processes.

K-CAP'01, October 22-23, 2001, Victoria, British Columbia, Canada.

KNOWLEDGE MANAGEMENT IN E-BUSINESS

In e-business, all business transactions and associated data are recorded. This is an important resource of knowledge about the business that can be used for decision making.
Knowledge can be made available using advanced information and communication technology as e-knowledge. It is suggested, that e-knowledge is a foundation for e-business [5]. E-knowledge is not only limited to descriptive knowledge or information, but also includes procedural and reasoning knowledge.

In the following, the term "knowledge" shall refer to knowledge that can be actively used in the generation of new information, that is procedural and reasoning knowledge. Descriptive knowledge, structured documents, associated data sets with meaning, and alike shall be referred to as "information".

The e-Knowledge Management Cycle

In the e-knowledge management cycle, business processes use information and knowledge to create new information [3]. Specific knowledge is extracted from this information by knowledge discovery and then utilised in the business transaction. The resulting knowledge management cycle is depicted in Fig. 1.

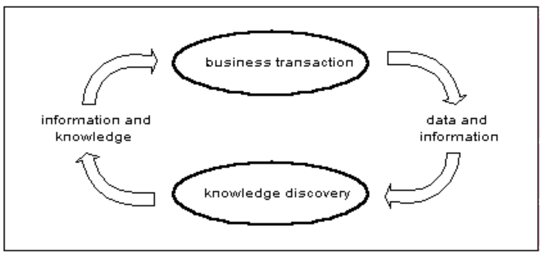

Figure 1: The e-knowledge management cycle

More precisely, all data and information associated to a business transaction are stored. This extends the more common case of only storing the customer and purchase data, thus allowing advanced analysis.

Information about products and production such as product and parts catalogues are included in the *business content* (the “what”). All kinds of knowledge and information for decision support in business is called *business intelligence* (the “how”). In general, business intelligence contains a combination of fixed business rules and adaptive business decision support.

The latter is extracted from stored data using intelligent data analysis. The analysis is applied on subjects of interest such as the group of customers, the set of products, the set of transactions, and so on. Results of analysis may also give raise to changing business policies. This yields flexibility to adapt to rapid changes in the business environment. Specifically, it can be used for knowledge-based marketing [12].

Information and Knowledge Flow

The e-knowledge management cycle is focussed on information and knowledge flow. In Fig. 2, a typical business process model of an e-business transaction is given. The upper part concerns the business processes, while the lower part depicts the specific information and knowledge required for the “e-business intelligence”. The arrows between business processes indicate flow control that may include information and knowledge flow. The lines with the bullet heads show the information and knowledge flow.

All data associated to the business transaction are stored. A transaction may be completed successfully, cancelled in one of the processes, or completed with complaints and possible settlements. This is in contrast to conventional business, where only the actual purchases are registered.

These data are useful for generating business reports, and more important, for extracting knowledge by *intelligent data analysis*. This provides an excellent view on the business environment. Conventionally, data analysis is only performed periodically to generate business reports. Beside these periodical runs, intelligent data analysis can be performed after each transaction, or if the situation gives rise to. There are several different analyses that involves information and knowledge flows:

- *Periodic reporting* may give rise to refine, update, or even change the associated knowledge.
- In each *transaction* as a check and detection of specific situations. Small updates can be implemented automatically.
- On *detection* of a specific situation. Especially if it is adverse, it may require more investigation, and as a result changes in policies.

The most important information in the *business transaction* is associated to the *quotation* as it progresses in time.

$$
\begin{array}{l}
request\ for\ quotation:\\
\quad (\ requested\ product\ specification,\ price\ range,\\
\qquad requested\ delivery,\ completed\ or\ cancelled\ )\\
quotation\ proposal:\\
\quad (\ proposed\ product\ specifications,\\
\qquad proposed\ price\ ranges,\\
\qquad expected\ delivery\ ranges,\\
\qquad completed\ or\ cancelled\ )\\
ordered\ quotation:\\
\quad (\ product\ specification,\ sales\ price,\\
\qquad planned\ delivery,\ completed\ or\ cancelled\ )\\
completed\ transaction:\\
\quad (\ product\ specification,\ sales\ price,\\
\qquad actual\ delivery,\ payment,\ settlement\ )
\end{array}
$$

The *specification and pricing* part controls the knowledge for the quotation and product specification. The product specification is is related to the production specification. The results of the intelligent data analysis are used to update and refine the required knowledge. Results may also give raise to changes in the production management.

Product Specification and Pricing

In e-business, the average customer should be able to specify the desired products online without the interference of a salesperson or an engineer. In addition, customers expect “online” speed. These requirements limit the complexity of the products. There are three categories of suitable tangible products:

- Off-the-shelf goods are products that do not need any processing before delivery (except packaging). These goods are fully specified by the catalogue of products, and only need intelligent searching. There are already quite a number of e-businesses of this type, such as booksellers.
- Composed products are assembled from ready parts on the customer’s specification. The parts are selected from catalogues, there are constraints on connections between parts. Therefore, examples and standard templates are important. These products are built-to-order, and there are some possibilities to change the configuration afterwards. There are a number of e-businesses, for instance PC shops.
- Flexible manufactured products are manufactured according to the customer’s specification using highly advanced multi-purpose production machines. Similar products can be manufactured in different ways, therefore the production specification is also needed. Because of its high flexibility, it is necessary to provide examples and templates. These products are customised built-to-order, and can not be changed after the production.

Catalogues of products are part of the *business content*.

Pricing of composed products and flexible manufactured products are elaborate. The price is composed of the estimated cost, additional margins, and profit [6]. Accurate calculation of the estimated costs can be fairly complicated, especially in flexible manufacturing [7]. Costing is usually based on *activity based costing*, thus attributing any expenses directly to the product. Activity based costing can be integrated in the enterprise system [13].

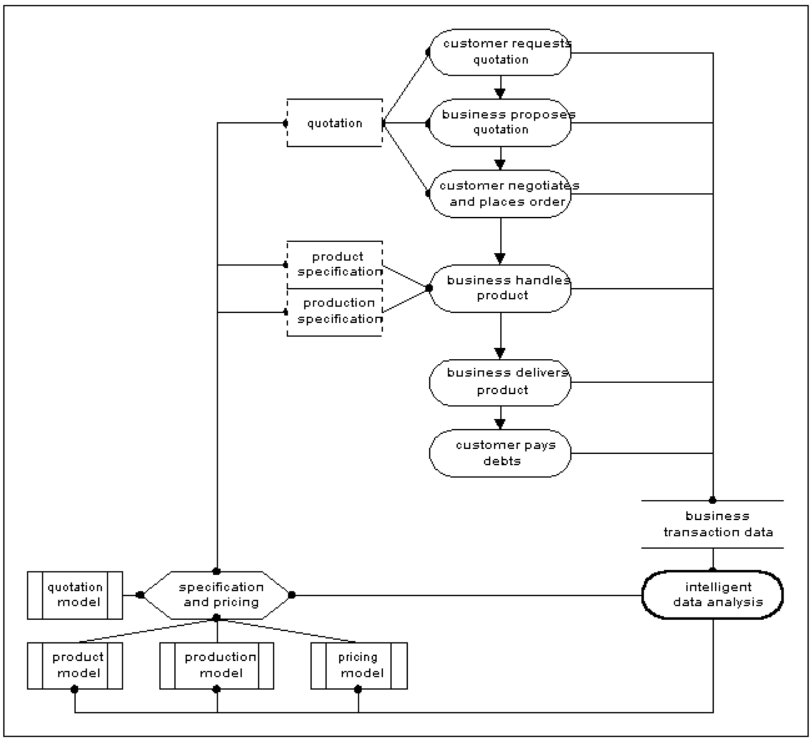

Figure 2: The information and knowledge flow in the e-knowledge management cycle

The following *pricing model* is proposed:

$$\begin{aligned}&calculated\ price = estimated\ costs + \\ &\quad risk\ with\ cost\ variances + \\ &\quad risk\ with\ estimation\ mistakes + \\ &\quad markup\ on\ materials + \\ &\quad markup\ on\ production\ costs + \\ &\quad profit\ margins\ ; \\ &estimated\ costs = estimated\ material\ costs + \\ &\quad estimated\ production\ costs + \\ &\quad estimated\ overhead\ costs\ .\end{aligned}$$

The final *sales price* is determined by the *pricing policy* that may depend on quality, quantity, and delivery; the customer, market situation, and other circumstances. For instance, "good customers" are rewarded by discounts, there may be a contract, potential customers can be attracted by special offers, special or rapid deliveries in a busy period are more expensive. Pricing can be adjusted by a simple decision rule, for example: if $customer\ credits > 10000,$ then $sales\ price = 90\%\ calculated\ price$. Decisions and the way regarding business is handled is *business intelligence*.

After each transaction, the realised costs can be computed and compared to the estimated costs. The $cost\ difference$ is used to improve the costing, especially for handling inaccuracies, and to recalculate the risks and margins.

$$\begin{aligned}&realised\ costs = material\ costs + \\ &\quad production\ costs + overhead\ costs\ ; \\ &cost\ difference = \\ &\quad estimated\ costs - realised\ costs\ ; \\ &profit = sales\ price - realised\ costs\ .\end{aligned}$$

The realised costs is the *price floor*, the lowest possible price without making losses. The highest price that the customer is willing to pay, the *price ceiling* is determined by the market.

AN INTELLIGENT INFORMATION SYSTEM FOR E-BUSINESS

An *intelligent information system* provides the usual *information system* functionality and *artificial intelligence* functionality in an integrated manner. It is a combination of an information system and an associated *expert system*. The intelligent information system for e-business should provide:

- intelligent business process support, especially in the online interaction;
- business content and business intelligence for adequate support of the quotation;
- advanced support of the e-knowledge management cycle.

The functionality correspond to those indicated in Fig. 2.

Intelligent Business Process Support

The major difference between conventional business and e-business is the online interaction with the customer, and that all information is recorded. Interaction is accomplished through an intelligent electronic form the *e-quotation*, that is generated according to a flexible *quotation model*. The quotation must be fully customised to the customer's needs, and at the same time transparent and clear in its use. To support online negotiations, the proposal should provide several product choices with proposed quantities, delivery times, and prices on a customer's request.

To the customer, the quotation contains a subset of the product specification that is understandable. It should also provide sufficient information on the products that meet the specification and further processing of the order. For further handling, the full product and production specification is needed. For convenience, frequent customers may have their own quotation and product templates. The business may also supply standard templates that may relate to the outcome of data analysis, for instance to promote new products.

The *specification and pricing* part is an intelligent part that provides expert system functionality for the quotation. The *product model* and *production model* are typically a combination of catalogues (business content) and business rules (business intelligence). The structure of a product model is basically hierarchical with interconnections and constraints [8]. Product choices are generated by matching the specified products to the production capabilities. The *pricing model* is generally highly adaptable to the business environment, such as changes in product costs, customer behaviour, and market demand.

The knowledge models consist of

- A qualitative definition part with a fixed structure that can only be changed by remodelling. In the product model, it is represented by an abstract hierarchical structure and interconnections of parts and materials. The production model is a sequence of manufacturing activities. The pricing model is determined by the formulas given earlier.
- A quantitative specification part that can change according to data. A product specification is a product model in which the parts and interconnections are filled out. The same applies to the production specification of this product. In pricing, the variables are replaced by values according to real data.
- An intelligent part that mainly contains decision support functionality, usually as logical rules and facts. In the quotation proposal, several options of similar product specifications are offered. The intelligence handles the matching of the product specification to available production machines in conjunction with additional business policies.

Intelligent Data Analysis

Intelligent data analysis [2] and data mining [4] has become an important issue because of the availability of large amounts of stored data. Data analysis is the process of transforming data into a meaningful form ("information"), generally different from the original format. Intelligent data analysis (IDA) is the analysis of complex data sets using artificial intelligence methods, thus extracting a knowledge model. The knowledge model is then used for further analysis and decision support.

From the viewpoint of application, data mining and IDA can be divided into four categories:

- *Summarisation* gives a compact description, generally by simple statistical analysis and tabulating the results such as mean, median, and standard deviations. Advanced multivariate visualisation at lower dimensions give an understandable overview of expected patterns and concepts. More advanced methods are the discovery of simple patterns that characterise the data as a small number of dominant concepts, relations, and deviations.
- *Dependence* describes significant deterministic or probabilistic dependencies between concepts. It is described in two levels: qualitative specifies which variables are dependent on each other, and quantitative specifies the strength of the dependence as a numerical value. Opposite to dependence is the detection of significant changes in the data from the previous or normative value. This includes the determination of deviations and outliers in a dependent group.
- *Grouping* divides into different homogeneous groups with clear concepts. The groups are fully determined by the concepts and the dependence on the concepts as variables. In *classification*, the groups are disjoint, while in *clustering* the groups may overlap. The number of clusters may be unknown (unsupervised), while the number of classes are fixed (supervised).
- *Prediction* is the estimation of an unknown or future values based on other known values, also referred to as forecasting. Prediction is the primary goal of IDA, on the other hand, the resulting knowledge model of data mining can be used for problem-solving. Grouping can be used for prediction by assigning the case to a group, and using the characteristics of the group to determine the unknown variables.

The main goals of data mining are *description* and *grouping*. Description is generally performed before any other analysis in conjunction with pre-analysis of the data. The primary goal of IDA is problem-solving, in particular *prediction*. In a data-driven situation, the actual IDA is preceded by data mining.

A number of important aspects of the business environment can be analysed:

- *Customer profile* by grouping according to characteristics such as purchases, products, and interaction behaviour. For instance, the groups of customers who visit and do not buy, visit and buy later, or visit and buy immediately can be visualised.

- *Product profile* by grouping and ranking according to popularity and associations to customer profiles. It is also interesting to know which products are purchased in combination.
- *Production method* with regard to efficiency and quality. This may lead to using specialised machines instead of multi-purpose machines for large amounts or special quality.
- *Pricing policy* associated to customer profile, product profile, and production method. Good customers may be rewarded by discounts, while popular products can be priced more accurately.
- *Sales forecast* for different types of products. This may lead to decisions on the product range.

Incorporating the e-Knowledge Management Cycle

Information systems commonly already incorporate small parts of the knowledge management cycle, as periodic reporting (summarisation). Other analyses may be periodical, but are usually invoked manually.

Fully automatic incorporation of the e-knowledge management cycle is not all that simple. In particular, control on updates of the knowledge models should be considered carefully as it may introduce inconsistencies. Changes to the quantitative part can either be implemented directly, or stored first and implemented after approval. Both the reason to, and the change should be verified and validated. In any case, deviations always need additional judgement to find out what the underlying reasons are.

It may thus be better to activate the cycle periodically, for instance in conjunction with the routine back-up. In fact, there are several knowledge management cycles simultaneously, each requires a different approach.

- *Periodic reporting* for customer reporting, product reporting, and production reporting tabulates characteristics of the subject in a period of time. Reports may give raise to business decisions.
- *Profiling* is determined from periodic reporting resulting in the customer profile, product profile, and production method by grouping. Profiles can be updated automatically by assigning each new case to a group. Failure can either mean that this case is a deviation, or that the groups are not appropriate and need to be renewed.
- The *transaction cycle* is activated at each transaction. It controls the most important information associated to the business transaction mentioned earlier and can therefore detect deviations.
- The *recalculation cycle* is invoked after each transaction and periodically. For pricing, it compares the calculated prices to the realised ones. It should signal on adverse situations, such as losses or other deviations. The result may be used to automatically update the variables of the pricing model.
- *Forecast* may be based on the temporal behaviour of the profiles. This may lead to decision on the product variety, production capacity and utilisation.

EXAMPLES

Because of its nature, it is expected that the proposed approach will be most beneficial for *flexible manufacturing*. There are several relevant flexible manufacturing industries: graphic arts, textile printing, plastics and multi-component moulding, packaging, cable industry, and metal industry. Although the above products seem very different from each other, the business characteristics are similar. All products are ordered to customer's specification, allow for mass customisation, are produced on demand by highly advanced machines, and have a short production life cycle. Graphic arts serves as a good example because it is widely available to average customers.

Varnavarni, a Graphic Arts Business

Graphic arts is an example of flexible manufacturing that has been around for a long time. Nowadays, the production machines are highly advanced and multi-purpose. The case presented here is based on PricingOnDemand [14], papers on this subject such as [9, 1], internal reports, examples of orders, and interviews with domain experts.

Varnavarni is a small e-business enterprise in quality printing of artworks such as posters, flyers, and postcards. All interactions with customers are fully online. The artworks are supplied online by the customer in one of the acceptable file formats.

The product characteristics for the quotation are:

- Image resolution: medium to high.
- Colours: black and white or colour in different qualities.
- Paper: a large variety of types, quality, thickness, and size.

The business has 3 types of production machines:

- Machine type A: 3-colour medium resolution, medium quality, high speed printer. There are 3 printers of type A: A1, A2, and A3.
- Machine type B: 3-colour high resolution high quality photo printer. There are 2 printers of type B: B1 and B2.
- Machine type C: 6-colour high resolution super high quality photo printer. There is 1 printer of type C.

A customer places a request for a black and white artwork in a limited number of high quality posters and a large number of medium quality flyers.

Request for quotation
Product specification
High quality
- Image resolution: high.
- Colours: black and white, high quality.
- Paper: glossy quality photo paper, size 600 mm x 900 mm.
- Amount: 13.
- Price range:

Standard
- Image resolution: medium.
- Colours: black and white, medium quality.
- Paper: fine quality, size 200 mm x 300 mm.
- Amount: 100.
- Price range: approximately 200.

Delivery: in 1 day.

Quotation proposal

High quality
- Image resolution: high.
- Colours: black and white, high quality.
- Paper: glossy quality photo paper, size 600 mm x 900 mm.
- choice 1 (type B): amount 13 price 335,
- choice 2 (type B): amount 15 price 355,
- choice 3 (type C): amount 13 price 400,
- choice 4 (type C): amount 15 price 425.

Standard
- Image resolution: medium.
- Colours: black and white, medium quality.
- Paper: fine quality, size 200 mm x 300 mm.
- choice 1 (type A): amount 100 price 185.
- Paper: super quality, size 200 mm x 300 mm.
- choice 2 (type A): amount 100 price 210.

Delivery: in 1 day. No additional fast delivery costs.

An approximate price range yields an interval of $\pm$ 10%. This customer has already purchased several times, therefore there are no additional fast delivery costs. The additional choices for the amounts of 15 instead of 13 is to promote the high quality products. Because the image is black and white, it does make sense to choose the 6-colour machine type C. Such irrelevant choices can be avoided by putting more intelligence in the product matching part.

Ordered quotation

High quality
- Image resolution: high.
- Colours: black and white, high quality.
- Paper: glossy quality photo paper, size 600 mm x 900 mm.
- choice 2 (type B): amount 15 price 355.

Standard
- Image resolution: medium.
- Colours: black and white, medium quality.
- Paper: super quality, size 200 mm x 300 mm.
- choice 2 (type A): amount 100 price 210.

Delivery: in 1 day. No additional fast delivery costs.

Examples of Knowledge Models

The products are printed graphics artworks on paper, and the production is the process of applying ink on paper according to the electronic source supplied by the customer. This is flexible manufacturing performed automatically by a single machine. The production method is fully determined by the machine by adjusting the controls at set up. If needed, paper, ink, and machine parts are replaced before adjusting the controls.

The production model is similar to the product model, the qualitative definition part contains the following items:

Product model	*Production model*
artwork in electronic form	
	production machine
paper type, quality, weight	paper type, quality, weight
paper size	paper size
ink colours, quality	ink colours, quality
image resolution	image resolution
	machine set up
	material waste
	machine operation

Paper size is a separate item, because it depends on the machine. The quantitative part gives the actual values in the model for further calculation. For example:
paper size: type = sheet; size = A3
paper size: type = roll; width = 600 mm, length = 9 m
machine operation: machine speed = 5000 standard imprints per hour.
Material waste is caused by misprints given as a percentage of a total of similar products for a machine type during a certain period.
material waste: if production machine = C; ink colours = full colour, high quality; then misprints = 10 %.

The choices of suitable production machines are determined by matching the items in the product model to the available production models.

The pricing model is determined by combining all items in the product model and production model. The estimated material costs is due to paper, ink, machine parts, and material waste. The estimated use of paper can be calculated accurately by matching the size of the product to the size of the paper in the machine.

The use of ink depends on the coverage of colours in the artwork and on the absorption of the paper. The latter can only be estimated from the technical specifications of the ink and paper manufacturers, and by experiment. The deviations of the real values from the initial specifications are unfortunately too significant to be included in the risk with estimation mistakes or markup on materials. These specifications merely give a lower bound.

Examples of the E-knowledge Management Cycle

In this section, a number of aspects common to this industry are mentioned. The application of the e-knowledge cycle, as well as the intelligent data analysis method and the required data are indicated.

The periodic product reporting revealed that there are approximately equal amounts of colour prints and black and white prints. It is therefore decided to use printer B1 for black and white prints, and printer B2 for colour prints. This decreases the machine set up time. Most reporting is generated by statistical summarisation on quotation data.

In the transaction cycle, it is found that an order has made losses. This could be corrected for by directly increasing the profit margin. But the reason for this were some unexpected problems during the set up which significantly had increased the set up time, and thus the job costs. In the intelligent data analysis, this is classified as an outlier.

The amount of misprints and the ink use have to be computed in each recalculation cycle. These analyses also require production management data such as job times, actual produced amounts, actual material use, etc.

In the forecast analysis, it is found that there is a decrease in cancelled quotations of high quality products. Forecast analysis on the grouping shows that this group is steadily growing. Forecast by itself generally only requires quotation data, but the determination of the associated capacities needs production management data.

CONCLUDING REMARKS

As illustrated in the examples, the e-knowledge cycle aids in the business decision taking, production management, and costs management. The e-business processes and the above features must be supported by an intelligent information system. The design of the intelligent information system requires thorough understanding of the specific industries and the associated production management.

Mostly, business transaction data are only retained for financial purposes, thus limited to completed transactions. Production management data and production data are usually registered separately in each production machine in different formats. Most of the data is not in a suitable format for analysis. This means, that the integration of data and interfaces to existing information systems and the production machines is a significant part of the design.

Further research will be dedicated on design and modelling. It is certainly a challenge to define generic knowledge models that can be utilised for different industries. In addition, the control of updates on these models according to the knowledge management cycle calls for thorough investigation as it may introduce inconsistencies.

ACKNOWLEDGEMENTS

The author is indebted to Huub van't Hart, Sylvia de Vries-Wennink, and Paul Gimberg, all with FrontEnd BV, for providing valuable domain expertise. Paul Gimberg has also given detailed comments. This research is jointly supported by the Centre for Telematics and Information Technology, the chair of Information Systems, Department of Computer Science, University of Twente, and FrontEnd BV.

REFERENCES

1. S. Almutawa and Y. Moon. The development of a connectionist expert system for compansation of color deviation in offset lithographic printing. *Artificial Intelligence in Engineering*, 13:427–434, 1999.
2. M. Berthold and D. Hand, editors. *Intelligent Data Analysis*. Springer, 1999.
3. U. Borghoff and R. Pareschi, editors. *Information Technology for Knowledge Management*. Springer, 1998.
4. U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy. *Advances in Knowledge Discovery and Data Mining*. AAAI Press / MIT Press, 1996.
5. C. Holsapple and M. Singh. Toward a unified view of electronic commerce, electronic business, and collaborative commerce: A knowledge management approach. *Knowledge and Process Management*, 37:151–164, 2000.
6. B. Kingsman and A. de Souza. A knowledge-based decision support system for cost estimation and pricing decisions in versatile manufacturing companies. *Int. J. Production Economics*, 53:119–139, 1997.
7. T. Koltai, S. Lozano, F. Guerrero, and L. Onieva. A flexible costing system for flexible manufacturing systems using activity based costing. *Int. J. Production Research*, 38(7):1615–1630, 2000.
8. T. Männistö , H. Peltonen, T. Soininen, and R. Sulonen. Multiple abstraction levels in modelling product structures. *Data & Knowledge Engineering*, 36:55–78, 2001.
9. K. Musselman. Complex scheduling of a printing process. *Computer & Industrial Engineering*, 39:273–291, 2001.
10. D. O'Leary. Enterprise knowledge management. *IEEE Computer*, 9:54–60, 1998.
11. D. O'Leary and R. Studer. Knowledge management: An interdisciplinary approach (editorial). *IEEE Intelligent Systems*, 16(1):24–25, 2001.
12. M. Shaw, C. Subramanian, G. Tan, and M. Welge. Knowledge management and data mining for marketing. *Decision Support Systems*, 31:127–137, 2001.
13. I. Tatsiopoulos and N. Panayiotou. The integration of activity based costing and enterprise modeling for reengineering purposes. *Int. J. Production Economics*, 66:33–44, 2000.
14. H. van 't Hart and S. de Vries-Wennink. Pricing on Demand. FrontEnd BV, www.frontend.nl, 2000.