An Incremental Approach to Maintenance: Its Influence on the Success of Expert System

Nur Azzah Abu Bakar
UUM College of Arts & Sciences
University Utara Malaysia
Sintok, Kedah, Malaysia
nurazzah@uum.edu.my

Fadzilah Siraj
UUM College of Arts & Sciences
University Utara Malaysia
Sintok, Kedah, Malaysia
fad173@uum.edu.my

***Abstract*— The difficulty of maintenance has been stated as among the factors contributing to the abandonment of many commercial expert systems. Ripple down rules approach to knowledge acquisition and representation has made maintenance of such system easier and less expensive. This paper presents the insight gained from four healthcare organizations which are currently using LabWizard™, an expert system which has been incorporated with ripple down rules technology. It is the findings from the work in progress to establish the factors leading to a long-term success of commercial expert systems. Qualitative data are gathered using case study approach, and analyzed using NVivo 8. The result shows that less difficult maintenance is one of the critical factors which can influence the success of expert system in routine use.**

Keywords-component: expert systems; knowledge acquisition; expert system maintenance; success factor.

I. INTRODUCTION

The term expert systems (ES) refers to an intelligent program which is designed to simulate the problem solving behavior of a human who is an expert in a narrow domain or discipline [1]. It is one of the earliest applications of Artificial Intelligence (AI) that achieves commercial viability. The system models two major traits of experts: (a) their expertise which is often referred to as domain knowledge and (b) their reasoning procedures. The basic idea behind ES is that, expertise is stored in the system's knowledge base to be called upon by users who wish to obtain specific advice or recommendations [2]. The system works by manipulating a declarative representation of the knowledge pertinent to a problem [3]. Like a human consultant, it gives recommendations and explains, if necessary, the logic behind its recommendations [4]. Such recommendations and explanation are useful for less experience users to allow them to make decisions which previously could only be made by human experts.

Besides, ES is also used to assist domain experts. In the case of R1/XCON [5], for example, the system was developed to assist computer configuration experts working at Digital Equipment Corporation (DEC). Due to the increasing volume of new computer system components and orders from clients, DEC experts were in real need for a tool to support them on a routine basis. In another example, pathologists working in clinical laboratories were put under pressure to deal with the high volume of test orders as this implies to the increasing number of interpretative reports to be generated. This can be the most challenging task for the pathologists as they need to process and recall large amount of information in order to make informed decisions within a restricted time period [6, 38]. Systems such as PUFF [7], PEIRS [8], HEPAXPERT-1 [9], TETANUS [10], VALAB [11, 12], REPCAT [13], CATIPO [14], VIE-PNN [15] and LIFECODE [16] are among the examples of ES which have been deployed in clinical settings. PEIRS, however, was abandoned four years after its inception while others are reported as either still in use or in unknown status. PUFF has been decommissioned but its knowledge base was then incorporated into Pulmonary Consult products. ES is beneficial to domain experts in a sense that it automates the most critical part of their jobs and so frees them from their routines. Organizations can also benefit from the existence of ES as they can make their organizational expertise available for other critical tasks in which ES is not yet available to assist. As witnessed in [11], 75% of the tedious, iterative and important work can be reduced and so allowed the experts to focus only on the reports that were rejected by the system.

It should be noted that, despite remarkable performance, only few ES succeeded in achieving long-term routine use or grow significantly in response to changes in the domain knowledge. Gill [18] surveyed 97 first wave ES and discovered only a third had continued to thrive. The failure-to-success ratio for ES was claimed as high as 10 to 1 in [20]. The technology was also made a slow progress in medicine [19] in which out of 98 publications on medical ES reviewed in [39], only 9 systems were reported as still in use. This is somewhat a surprising fact as medicine is one of the earliest application domains for ES, and medical tasks are knowledge intensive tasks which generally are good candidates for ES.

The myriad of literature on ES focuses on various technical aspects of the technology. There are also plenty of published papers reporting ES prototypes in various application domains. To date, not much attention has been given to the deployed systems to see how far they have succeeded once installed at the organizations. Yoon et al. [37] draw attention to a serious lack of empirical evidence associated to success or failure of ES. Hence, this paper

978-0-7695-4262-1/10 $26.00 © 2010 IEEE
DOI 10.1109/CIMSiM.2010.91

presents some real evidences from practical experiences which are obtained from an investigation on a number of clinical pathology laboratories which are using ES in their daily routines.

II. MAINTENANCE ISSUES IN ES

Maintenance is referred to as a process of acquiring and incorporating knowledge from one or more knowledge sources into the knowledge base. ES relies heavily on its knowledge content. Due to the fact that knowledge is dynamic and changes over time, every ES requires frequent maintenance so as to ensure it contains only up-to-date knowledge and not quickly becoming obsolete. As for example, half of the thousands of rules in XCON's knowledge base are changed every year to adapt to rapid changes in computer components [36]. This is important to ensure the reliability of its recommendations or advice. This is extremely important if ES is used in medical domain as the system poses risks for patient safety by system malfunctioning or misuse. Quality and safety are among the issues that always receive a great concern in medicine [14, 17]. Therefore, ES must provide the medical users with accurate output to avoid any ethical or legal consequences.

When ES first came to prominence, there were claims that it lent itself to easy verification, extension and maintenance [3]. In contrast, these proved to be the difficult tasks to accomplish. Maintenance is difficult for various reasons. To name a few, the nature of the representation techniques used which did not actually lend themselves to traditional maintenance techniques; the frequency of changes in domain knowledge; the lack of supporting methodologies, techniques and tools; and the unfamiliarity of the members of maintenance team with the development techniques and methods used, particularly when they are not the members of the original development team. Human factors, such as skepticism, uncooperative or unwilling experts, as well as unrealistic understanding and expectation towards ES technology can also exacerbate the situation. Furthermore, when multiple experts involved in maintenance, they tend to disagree between each other on the knowledge to be used.

The early approach to maintenance is knowledge engineers work closely with the domain experts to acquire their knowledge using various knowledge acquisition techniques. Experts are required to explain, normally through a retrospective approach, how they reached a specific recommendation. The acquired knowledge is then analyzed and coded into the system's knowledge base in the form of rules, frames or other knowledge representation schemes. Studies in user-centered philosophy, however, claimed that experts never explain how they reached a specific conclusion. Rather, they justify why their conclusions are correct [26, 27]. This philosophy serves as a basic principle underlying the development of Ripple-Down Rules (RDR), an incremental knowledge acquisition technique which is also a knowledge-based system construction methodology.

Knowledge in RDR system is acquired incrementally only in a context it is used [25]. As a result, it simplifies the processes of acquiring and incorporating knowledge up to a stage where the domain experts can take responsibility to maintain the system on their own. Details of how RDR system works can be found in (e.g., [26-28]). The first commercial application of RDR is PEIRS, a large medical ES for the interpretation of chemical pathology reports. An initial validation with PEIRS proved RDR efficiency and performance in real practical problem.

III. LITERATURE REVIEW

Prior researches (e.g., [19-25]) provide some evidences of the relationship between the various issues related to maintenance and the abandonment of many commercial ES. Besides, many other factors were also discussed in the literature, as synthesized and shown in Fig. 1. Taking PEIRS as an example, the system fell into disuse at Sydney's St. Vincent Hospital when the hospital set up its new hospital information system (HIS). Less priority to integrate PEIRS with the newly setup HIS was said to be the main reason for discontinuation of PEIRS. Nevertheless, our discussion in this paper focuses only on maintenance as well as on other technological and non-technological factors which are related to it.

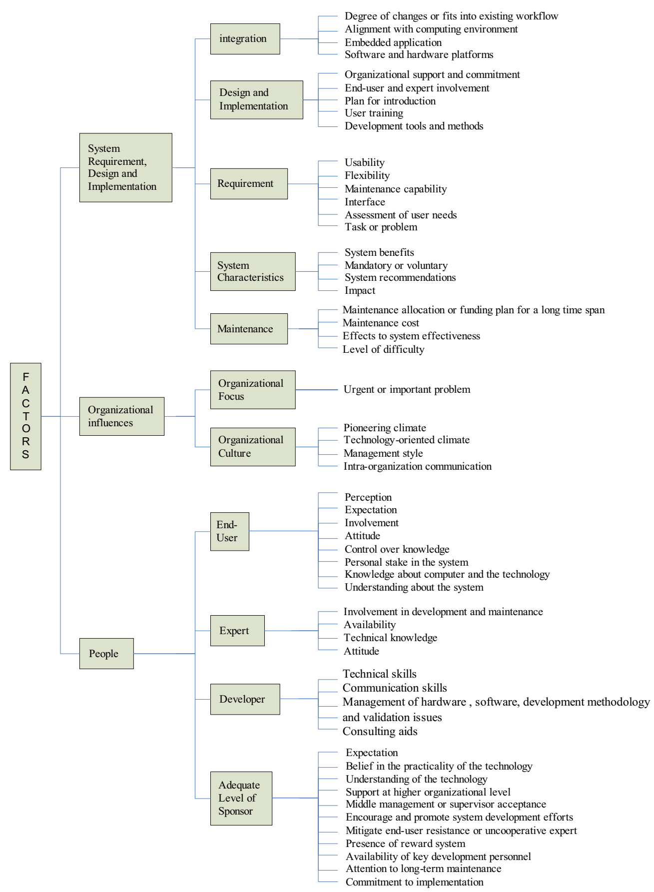

Figure 1. The general framework

The figure serves as a conceptual framework for our study, aiming to provide general understanding of what had influenced people to continue or discontinue using ES. It should be noted that the figure is neither complete nor comprehensive. We strongly believe there are still a lot of practical experiences which had never been reported anywhere. The factors listed are also not bound to the definition of success described in the next section, and are not domain specific. Other than maintenance, factors such as domain expert involvement and system explanation are also pertinent to the unique characteristics of ES. Hence they became the focus of this study as a whole. This was based on our discussion with the field expert, and also on the insight gained from the literature. Coenen and Bench-Capon [3], for example, argued that "for ES to be commercially viable it must be able to respond to the changes in the domain knowledge on which it is based". In other words, it needs to be maintained.

IV. METHODOLOGY

This study employs a qualitative research method. Specifically, case study design is used as this study is an exploratory research aiming to investigate the phenomena within its real-life context. The objective is to investigate why some ES are long lasting while many others failed shortly after being deployed. No established theory is applicable to this question, thus the use of case study is deemed to be appropriate. Kitchenham et al. [31] argued that case study is as an important method for industrial evaluation of software technology. It has also been recognized as a formal research method in software engineering, and is becoming popular within the field. Study is conducted in six phases according to the *Guidelines for Industrially-Based Multiple Case Studies in Software Engineering* [32] comprises of pre-planning, administration, planning, data collection, data analysis and documentation phase. Due to the fact that not much information can be found on successful ES, case study is the most appropriate method to be used as it allows an in-depth investigation on each case. A number of previous researches used this method when investigating an impact of intelligent systems on specific organizations [e.g. (33-35)]. A survey method requires a significant volume of data which are inaccessible in this case, thus it is not appropriate for this research.

A. Success Definition

The term *success* used in this paper is defined based on three criteria: (a) the longevity of the system in actual operational use; (b) the extent of use; (c) the growth of the system's knowledge base. The period of five years is argued in [18] as enough to measure user penetration and longevity. The extent of use, on the other hand, reflects the importance of the system to users which may foster its continuous use. The third success definition, i.e. the growth in the knowledge base, reflects the capability of the system to be maintained to deal with changes in the domain knowledge.

B. The Investigated System

The system that is being the focused in this research is LabWizard, a commercial product of Pacific Knowledge System (PKS) which is in use in more that 30% of all Australian pathology laboratories [29, 30]. There are two versions of LabWizard namely LabWizard Business Auditing System and LabWizard Clinical Reporting System.

LabWizard was chosen for the reason that it fulfills all the criteria for a successful ES as defined in section IV.A. In contrast to PEIRS which was also incorporated with RDR, LabWizard is being in routine use since its inception, and serves as an integral part of the organization's day to day operation. In addition, some of the laboratories using LabWizard managed to expand the system's knowledge base to contain more than 10,000 rules and provide about 1.6M interpretations per month [30].

C. Data Collection

The data reported in this paper are obtained from healthcare organizations in New South Wales (NSW), Australia Capital Territory (ACT) and Queensland (QLD), Australia. This stage involved four clinical laboratories including both hospital-based laboratory and private laboratory.

As laboratory ES are embedded within the process of care, they usually do not intrude into clinical practice. Clinicians working with patients do not need to interact with ES but benefit from the system's report as it contains a diagnostic hypothesis for consideration. System also does not remove clinician's responsibility for information gathering, examination, assessment and treatment. For the pathologists, the system cuts down the workload of generating reports without removing the need to check and correct them. By investigating laboratory ES some issues with regard to poor integration with the existing workflow can be eliminated.

The primary source of data for this study came from interviews which were conducted with five interviewees at their sites. Those who were selected for interviewing were people who either have direct interaction with LabWizard or possess understanding of the issues surrounding its deployment, at least at their own site. Some data were also obtained from secondary sources such as published research papers, industry reports and documentation provided by the interviewees as well as through observation. Open-ended interviews were used to give more freedom to the interviewees to express their opinions.

D. Data Management and Ananlysis

A qualitative data analysis software NVivo8 is used as data management and analysis tool. The body of literature is also useful and so serves as the background data and created as part of the main project. The interview transcripts were imported into the software and examined for themes and coded accordingly. The preliminary model, which was developed, based on the background data was continuously refined in response to the emerging themes from the case studies.

V. FINDINGS

A. *Business Auditing System*

The system, also known as Auditor, was developed specifically to facilitate the detection of data entry errors and billing anomalies, hence serving as a front-end system for the user. When the laboratory information system (LIS) receives a request for test order, LabWizard picks up the details such as details of the current request, billing, confidentiality codes, patient data, patient visit data as well as details of the ordering clinicians or general practitioners (GP). The system uses the rules contained in its knowledge bases to assess the risk of error, and flags the entry if there is any. Thus, it reduces the number of entries to be reviewed and corrected, if necessary, by the senior operational staff .

Three out of the five respondents are the users of this system. They do not possess good technical background in computing. Rather, they have background knowledge of the laboratory system such as the business rules, billing rules, protocols and standards that they have to meet. The general factors which led to the continuous use of the system include the benefits which can quickly be realized, the user-friendly interface and the ease of use. Besides the reduced number of errors, time saving is another benefit most appreciated by the users. Prior to LabWizard, users perfomed their task in a very manual fashion; they ran an SQL through a database each night which pull out every data entry that was made in the previous 24 hours and had it formatted into Excel spreadsheet so that it can be sorted and filtered. The process normally took them between 4 to 5 hours a day everyday, as compared to LabWizard which only took about an hour a day. One of the respondents described the system as *"something we have been waiting for for a long time"*. The system has helped them to efficiently deal with the increasing volume of request for test order. At the same time, it enabled them to protect their patients' privacy, for example, by ensuring that the report is sent back to the right GP.

With regard to maintenance, easy, rapid and cost-effective maintenance have contributed to the system's success. The findings confirmed that the incremental techniques used in the system, which is actually based upon RDR technology, helps them to maintain the system effectively without having to learn the tips and tricks of programming. The time required for adding a simple rule varies between the respondents in a range between 2 to 5 minutes. This is far better than the standard performance in which only 1 or 2 rules can be added in a day. The respondents found no difficulty to identify what to be added into the knowledge base whenever there is a need, as the system provides some means of triggering them with the required knowledge through the use of cornerstone cases. In addition, the system also facilitates users to recognize the situation when new rules must be added.

As LabWizard is able to automatically patch the new rules, users who maintain it are not required to know the structure of the existing knowledge base and so eliminate part of the difficulties of maintenance. Continuous support from PKS, in which the respondents claimed as always available for help, has also been acknowledged as one of the factor for the respondents to continue maintaining the system, especially when more complex rules are to be added into the knowledge base.

B. *Clinical Reporting System*

The system provides an expertise for a routine work that runs silently in the background of the laboratories. It facilitates the provision of highly specific interpretive comments for laboratory results, as well as for managing workflows within the laboratory. The pathologists using this system can create and maintain their own knowledge bases of clinically accurate and patient-specific comments and recommendations that are applied automatically to the test results.

Correctly recognizing and interpreting pathological findings is critical; failure to do so can be fatal for the patients and thus not comply with the guideline for medical diagnostic. In general, the factors leading to success of this clinical system are found similar to those contributing to success of the Business Auditing System. This could probably due to the fact that the same technology is used as a basic principle underlying both systems. One of the pathologists, however, claimed that he is glued to the system because he is ensured that he reserved the rights for his expertise which is now stored in the system's knowledge base.

This system is maintained by the pathologists, who claimed that the process is easy and can be done quickly even though they are not equipped with knowledge engineering skills. They in fact can start with either an empty or seed knowledge base and quickly build on the knowledge bases to provide the specific comments as they require. As a result of doing the maintenance part by themselves, they feel ownership, and have control over the knowledge stored in the system. Hence, this promotes them to continue using and maintaining it as they go on with their routine work. A comprehensive interpretation report can then be generated from such a comprehensive knowledge base, but the system still needs to be well integrated with the existing LIS to get details of patients' historical data.

VI. CONCLUSION

Two systems that are investigated in this study represent instances of error detection ES and an interpretative ES. Both were built upon an incremental knowledge acquisition technique which allows the system to be easily maintained when needed. Recent literature has called for a better understanding of factors that predict success of systems supporting clinical decision making. Hence, findings from this study is significant to shape a better understanding of how maintenance should be done in order to promote continuous use of ES, as well as growth in its knowledge base.

One of the challenges in conducting this research is the lack of published practical evidences from the organizations. Very few publication focuses on deployment aspect of ES; if there is any, it is in a form of an anecdote which does not provide enough information for the researcher to better understand the phenomena. Future works in this area is still needed to improve our understanding on the issues surrounding the deployment of ES.

ACKNOWLEDGMENT

I thank Prof. Dr. June Verner and Emeritus Prof. Dr. Paul Compton of School of Computer Science and Engineering, The University of New South Wales for their opinions and continuous support for this work. I also thank Pacific Knowledge System (PKS), especially Dr. Lindsay Peters for granting me permission to conduct this study with PKS clients and for helping to initiate a network with them. I finally thank all the respondents for giving some of their precious time to share their experiences with LabWizard.

REFERENCES

[1] J. Durkin. Expert Systems: Design and Development. New Jersey: Mcmillan, 1994.

[2] S. Liao, "Expert systems methodologies and applications—a decade review from 1995 to 2004," Expert Systems with Applications, vol. 28, 2005, pp. 93–103.

[3] F. Coenen and T. Bench-Capon. Maintenance of Knowledge-Based Systems. London: Academic Press Ltd, 1993.

[4] E. Turban and J. E. Aronson. Decision Support Systems and Intelligent Systems, 6th ed. Prentice-Hall International, 2001.

[5] J. McDermott, "R1/XCON at age 12: lessons learned from an elementary school achiever," Artificial Intelligence, vol. 59, 1993, pp. 241-247.

[6] S. Williamson, "Pathology information system," Today's Life Science, 2001, pp. 28-29.

[7] J. S. Aikins, J. C. Kunz, E. H. Shortliffe and R. J. Fallat, "PUFF: an expert system for interpretation of pulmunory function data,".

[8] G. Edwards, P. Compton, R. Malor, A. Srinivasan and L. Lazarus, "PEIRS: a pathologist-maintained expert system for the interpretation of chemical pathology reports," Pathology, vol. 25, 1993, pp. 27-34.

[9] K. P. Adlassnig and W. Horak, "Development and retrospective evaluation of Hepaxpert-1: a routinely used expert system for interpretitive analysis of hepatitis A and B serologic findings," Artificial Intelligence in Medicine, vol. 7, 1995, pp. 1-24.

[10] J. P. Schroder, W. D. Kuhlmann and Chr. Trendelenburg, "Knowledge-based approach to clinical decision-support system with an application in tetanus serology," Clinica Chimica Acta, vol. 222, 1993, pp. 79-83.

[11] P. Vladiguie, E. Rogari and H. Philippe, "VALAB: expert system for validation of biomedical data," Clin Chem, vol. 38, 1992, pp. 83-87.

[12] P. Vladiguie, E. Rogari, J. X. Corberand and B. Boneu, "The performance of the expert system VALAB revisited: an evaluation after 5 years," Eur J Clin Chem Clin Biochem, vol. 34, 1996, pp. 371-376.

[13] G. A. Smythe and C. M. Drew, "REPCAT: desktop expert system for interpreting and validating laboratory data for phechromocytoma diagnosis with the database application Omnis 7," Clin Chem, vol. 43, 1997, pp. 134-140.

[14] P. Knaup, T. Wiedemann, A. Bachert, U. Creutzig, R. Haux and F. Schilling, "Efficiency and safety of chemotherapy plans for children: CATIPO—a nationwide approach," Artificial Intelligence in Medicine, vol. 24, 2002, pp. 229-242.

[15] W. Horn, C. Popow, S. Miksch, L. Kirchner and A. Seyfang, "Development and evaluation of VIE-PNN, a knowledge-based system for calculating the parenteral nutrition of newborn infants," Artificial Intelligence in Medicine, vol. 24, 2002, pp. 217-228.

[16] D. T. Heinze et al., "LIFECODE: a deployed application for automated medical coding," AI Magazine, Summer 2001, pp. 76-88.

[17] J. Fox and R. Thomson. (2006). Issues in the use of clinical decision support systems: efficacy and safety, ethics and liability. Advanced Computation Laboratory. [Online]. Available: http://www.acl.icnet.uk/lab/PUBLICATIONS/ms400.doc.

[18] T. G. Gill, "Early expert systems: where are they now?" MIS Quarterly, vol. 19, no. 1, 1995, pp. 51-80.

[19] H. Heathfield, "The rise and "fall" of expert system in medicine," Expert Systems, vol. 6, no. 3, 1999, pp. 183-188.

[20] J. A. Oravec and L. Travis, "If we could do it over, we'd … learning from less-than-successful expert system projects," Journal of System and Software, vol. 19, 1992, pp. 113-122.

[21] G. H. Subramaniam, G. J. Yaverbaum and S. J. Bandit, "An empirical evaluation of factors influencing expert system effectiveness," Journal of System and Software, vol. 38, 1997, pp. 25-261.

[22] R. P. Will, M. K. McQuaig and D. E. Hardaway, "Identifying long-term success issues of expert systems," Expert Systems With Applications, vol. 7, no. 2, 1994, pp. 273-279.

[23] C. K. Tyran and J. F. George, "The implementation of expert systems: a survey of successful implementation," ACM SIGMIS Data Base, vol. 24, no. 1, 1993, pp. 5-15.

[24] T. Wetter, "Lessons learnt from bringing knowledge-based decision support into routine use," Artificial Intelligence in Medicine, vol. 24, 2002, pp. 195-203.

[25] P. Compton and R. Jansen, "A philosophical basis for knowledge acquisition," Knowledge Acquisition, vol. 2, 1990, pp. 241-257.

[26] P. Compton, L. Peters, G. Edwards and T. G. Lavers, "Experience with Ripple-Down Rules," Knowledge-Based System Journal, 2006.

[27] P. Compton, B. Kang, P. Preston and M. Mulholland, "Knowledge acquision without analysis," European Knowledge Acquisition Workshop, 1993.

[28] B. Kang, "thesis title," UNSW Thesis

[29] G. A. Edwards, "Is LabWizard a logical extension of the electronic patient record?," Ned Tijdschr Klin Chem Labgeneesk, vol. 29, 2004, pp. 230-231.

[30] BETA 2010, http://beta.csesoc.unsw.edu.au/2010/05/the-life-of-paul-compton/

[31] B. A. Kitchenham, L. M Pickard and S. L. Pfleeger, "Case studies for method and tool evaluation," IEEE Software, vol. 12, no. 4, July 1995, pp. 52-62.

[32] J. M. Verner, J. Sampson, V. Tosic, N. A. Abu Bakar and B. A. Kitchenham. "Guidelines for industrially-based multiple case studies in software engineering," in Proceedings of the 2nd International Conference on Research Challenges in Information Science, Fes, Morocco, 2009.

[33] A. A. Baldwin-Morgan, "The impact of expert systems on auditing firms: evidence from a case study," International Journal of Applied Expert System, vol. 2, no. 3, 1994.

[34] J. Trewin, "The impact of the introduction of an expert systems on a professional accounting organization," unpublished.

[35] S. Zawa, "Case study of a knowledge-based system for consumer lending,".

[36] E. Soloway, J. Bachent and K. Jensen, “Assessing the maintainability of XCON-in-RIME: coping with the problems of a very large rule base,” in AAAI’87, 1987, pp. 824-829.

[37] Y. Yoon, T. Guimaraes and Q. O’Neil, “Exploring the factors associated with expert system success,” MIS Quarterly, 1995.