Dr. Sanaa Abdul Kariv

Sohar University
Faculty of Business

The 2nd International Conference on Applied Business and Economics - (2nd ICABE2010)

December 14-15, 2010

icabe_su@yahoo.com

Theme: INNOVATIVE STRATEGIES FOR SUSTAINABLE DEVELOPMENT

Proceedings

Venue: *Sohar University, Sohar, Sultanate of Oman*

Sponsors: *Sohar University*

Organiser: *Faculty of Business, Sohar University*

P. O. Box 44, PC 311, Sohar, Sultanate of Oman
Tel.: 0096826720101, Fax: 0096826720102
www.soharuni.edu.om

Application of Expert System Approach in Designing Knowledge Engineering system

Dr. Alkhanak, Sanaa
University of Malaya- Malaysia
sanalkhanak@yahoo.com

1. INTRODUCTION

The increasing use of Information Technology, telecommunications and networks, rely on digital systems to obtain the data and information, and knowledge, process, manage, transform, store, disseminate and sharing within and outside the organization with the achievement of the possibility of retrieval and update this information, all in turn led to help organizations achieve their goals in obtaining a competitive advantage enabling it to go through the global market. On the other hand, Knowledge plays a vital role for all acts of contemporary organizations, and benefit from developments in Information Technology systems and artificial intelligence, to capture the knowledge of human experts and make it accessible to everyone. This paper presents an attempt to develop a system for "knowledge engineering" cares about how organizations can extract its knowledge, especially implicit ones, and refer to it when was needed. Therefore, the **research question** arises, what are the relations between the system factors, and how to graduate the importance of system factors? Nonetheless, the **research objectives** is to take advantages of applications of artificial intelligence in suggesting a design for knowledge engineering system that simulates human intelligence in the storage and retrieval of knowledge, this design depends on a set of assumptions, components and processes. Moreover, in order to answer the questions raised in the research problem was formulated two main **hypotheses**: first: There are significant correlations (0.50) between the factors of suggested knowledge engineering system. Second: These factors are at different levels of important. In addition, to test hypotheses, a **questionnaire** was developed and distributed to a random sample of a group of computer science professors some universities and the correlation coefficient and factor analysis was used to prove the two hypotheses. Finally, the presence of different levels of knowledge (information, practice, experience) of sample members in regard to the system dimensions which reflects the possibility of its application. All the variables are important enough to be included as factors of the system.

2. LITERATURE REVIEW

Along with the previous studies that have dealt with different aspects of Knowledge Engineering are the study series that were done by Zhang et al. (2009) propose an expert system which is based on Knowledge Engineering named DDES - Dental Diagnostic Expert System. Besides, a hybrid knowledge acquisition strategy is proposed to efficiently elicit the knowledge from experienced dentists and construct the medical ontology. The system fulfills the storage of dental photos and X-ray images, typical cases, patients' information and concerned data. It enables the complex query for combined conditions and the display of requisite results by interactive interface between user and computer [1]. **Studer et.al (2000)** formulate a conceptual model for the knowledge engineering at the present time and the future, the study also investigate the importance of knowledge in solving the problems facing contemporary organizations[2]. (Swanson & Collaboratory, 1999) described how the Hidden public knowledge could be presented through the use of data and the design of different analytical software in an effort to invent suggested correlations of mathematical scenarios that aim at revealing the Hidden Knowledge and how to put it in use, furthermore [3], (Rivtee and Ktine, 1999) study has focused on Patents and Intellectual Properties as good examples of that [4]. Studer **et al. (1997) described in their research** three modeling frameworks: Common KADS, MIKE and PROTÉGÉ-II. This description is supplemented by discussing some important methodological developments in more detail: specification languages for knowledge-based systems, problem-solving methods and ontologies. In addition the research outlines the relationship of Knowledge Engineering to Software Engineering, Information Integration and Knowledge Management [5].

3. CONCEPTUAL FRAMING OF THE PROPOSED KNOWLEDGE ENGINEERING SYSTEM:

Knowledge Engineering is the use of practical means for the acquisition of knowledge from its sources and then represents it in the form of significance networks or patterns of grouped concepts that are stored in the knowledge base to gain some of the capabilities of the human mind mechanism. The intended term of Knowledge Engineering is the measure of reducing the vast amount of knowledge to a tight set of "facts and rules" in order to be stored on the knowledge base of the Organization (6), also identified as the process of developing knowledge based systems in any field [7]. For the purpose of developing a conceptual framework for knowledge engineering System, and in order to understand all aspects of this system, we need to use the tools of artificial intelligence to design an based on a set of assumptions, and expert System for knowledge engineering based on a set of assumptions, and components (physical and people) and processes outlined. Figure (1) shows the **components and operations of the Knowledge Engineering in proposed system**

1) THE SYSTEM ASSUMPTIONS:

- Knowledge Engineering deals with the acquisition of knowledge and its representation, to ensure its validity and inference, and finally its clarification and maintenance.
- The Knowledge Engineering System is a sub-system of Artificial Intelligence.
- The system consists of a set of Physical Components (hardware & software) and human Components (Knowledge engineer & expert), works through a series of operations.

2) THE COMPONENTS OF THE KNOWLEDGE ENGINEERING SYSTEM:

Knowledge Engineering System consists of two main components: Physical Components (hardware & software) and human Components (Knowledge engineer & expert) that work together to implement the system and processes as follows:

- **THE PHYSICAL COMPONENTS OF THE SYSTEM**

It presents the physical component which is used to implement operations of the system, it consists Knowledge Base, Knowledge Acquisition Facility, Inferences engine, Explanation Facility, User interface, The Development engine, Black board, Knowledge Refining System.

Knowledge Base: The repository of knowledge accumulated that can be used to support the end user or to support the complex decision-making (8). The knowledge in the Knowledge Base has the forms of rules, facts, experiences, and beliefs.

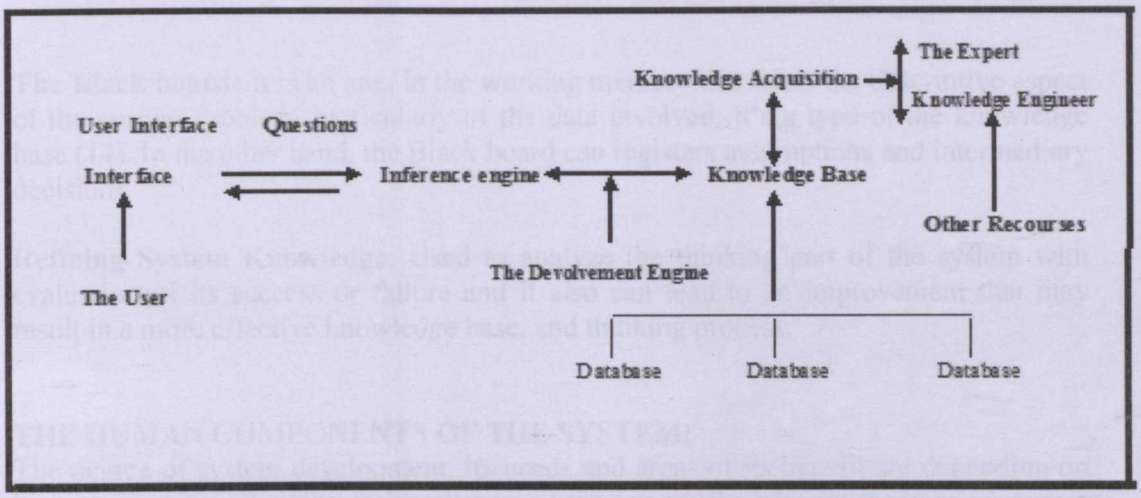

Figure 1 components and operations of the Knowledge Engineering
Source: The researcher

Knowledge Acquisition Facility: A tool to acquire knowledge from human experts(9), which is also the tool that allow the system to acquire more information about the problem area directly from the expert, or from other sources such as libraries and databases(10).

Inferences Engine: It represents the reasoning element in the system, the Inference Engine helps in "reasoning" to solve the problems addressed in the system (11) which operates on the basis of facts and rules, and act as the executive factor of the system that drives the system to solve problems as well as to modify the existing rules in the knowledge base via the addition or the modification of knowledge that is related to those problems. In addition, there are several methods to solve the problems of logical reasoning depending on the particular circumstances of the knowledge base rules, and they include:

- **The Forward Chaining.**
- **The Backward Chaining.**

Explanation Facility: It's the tool that allows the user to understand the logic and the reasoning behind the conclusions or results provided by the system, through explaining the facts and the rules that are used to reach the results (12). Moreover the explanation facility can provide the user with answers to the questions (Why) and (How).

User Interface: User Interfaces is designed according to the needs and requirements of the user in a way to facilitate the interaction between the system and the user (13).

The Development engine: Knowledge base analysis can be done through the examination of some examples that are designed to be processed by the system; if there are some logical conflicts to appear, and then the knowledge engineer would use the Development Engine "to correct the knowledge base".

The Black board: It is an area in the working memory that takes the descriptive aspect of the current problem, particularly of the data involved, it's a type of the knowledge base (14). In the other hand, the Black board can registers assumptions and intermediary decisions.

Refining System Knowledge: Used to analyze the thinking part of the system with evaluation of its success or failure and it also can lead to an improvement that may result in a more effective knowledge base, and thinking process.

THE HUMAN COMPONENTS OF THE SYSTEM:

The degree of system development, its needs and areas of its benefit are depending on human resources. Ranging from reliance on the expert, the user (in the case of system adoption of the rules of expert introduction of knowledge to the system), and the knowledge engineer interference as a knowledge programmer. This indicates the possibility of depending on the Knowledge Engineer, the Expert and the User together as the main human resources in designing this system.

Knowledge Engineer: He is the person who is responsible of the architecture of knowledge based systems; he creates and designs models as well as structuring the knowledge experiences of the expert (15). Besides that, knowledge engineer responsibilities are to guide the expert during the knowledge creation process, correcting the knowledge base, managing the knowledge encryption tools, as well as correcting the encoded knowledge base.

The Expert: The Expert is one of the knowledge resources that feed the system, but the knowledge possessed by the expert is complex in relation to the rest of the sources, and this leads to cause longer procedures for the processing of that knowledge to system. The expert's insight about the problem nature, its theme and complexity are the important parameters in choosing the proper Expert for the system.

The User (The beneficiary): Provision of the advice and counseling to the beneficiary user of the system is regarded as the core purpose for designing the Knowledge System. The system user can be a student at a university, an engineer working in a factory, or sales operative in a marketing company.

3) **THE KNOWLEDGE ENGINEERING OPERATION**

The system works through a series of operations:

KNOWLEDGE ACQUISITION

The basic model of acquiring knowledge is through the intermediary role of the Knowledge Engineer between the Expert and the Knowledge Base of the organization in a process of prevailing the knowledge from the Expert, encoding it and refining it for the purpose of establishing a Knowledge base. It can be summarized in the following stages: The knowledge engineer does interviews with the expert to acquire his knowledge, encoding the knowledge for the purpose of integration in the Knowledge Base, and the use of Shells in the Knowledge Base to get a precise indication of the knowledge status (16).

VALIDATION & VERIFICATION

It is a type of evaluation that seeks to establish a true system, a system that works at an acceptable level of accuracy and validation of the application system specifications, such as verification of accuracy and completeness, width, depth, sophistication and realism.

KNOWLEDGE REPRESENTATION:

It is the organizing of the acquired knowledge that was and its incorporation in the knowledge base. There are different methods of integrating the information and Knowledge in the knowledge base. The main basic methods of representation are, by logic, semantic arrays networks, multiple knowledge representation, and uncertainty representation.

INFERENCE PROCESSES:

Inference Processes being done through an algorithmic program that controls the process of reflection, which represents the intelligent and cognitive part of the system. Then the conclusion is reached depending on the method by which the knowledge has been represented in the Knowledge Base like: through forward and backward rules inference, inference hierarchy, and case-oriented reasoning.

KNOWLEDGE EXPLANATION:

The system responsibility is to explain the reached conclusions through the illustration tool. There are several purposes for this explanation function, among the most important ones are: The clarification of assumptions from the operations of the system, for both; the user and the establisher.

DEVELOPMENT OF THE SYSTEM

System development is achieved through the process of acquiring knowledge which, in fact, is carried on between the expert and the user as follows: the expert is elicits required data from the user, then he develops one model at least according to data obtained and generates the consultation (based on the applicable formula) and after that he will feed it to the user. The Customer, on the other side, accepts the advice or makes another questioning. (These questions lead to cycling in elicitation, modeling, design and questioning).

4) HOW DOES THE PROPOSED KNOWLEDGE ENGINEERING SYSTEM OPERATE?

The following Figure shows the work-flow of the proposed **knowledge engineering system**, to achieve its objectives, through using the system **physical and human components** according to the **assumptions** that have been set, The **Knowledge Engineer** acquire the documented knowledge as well as the undocumented one (that is resent in the Expert's mind) through the use of **acquisition tool** that allows the establishment of **The Knowledge Based** through the input **rules and facts** that have been acquired, after the **knowledge engineer had** verified the level of reliability and credibility (through a number of standards that are established for this purpose). Since the main goal of the system is to carry out the conclusions, decisions and solve problems to become ready for use by the end user (**beneficiary**) of the system, all of that can be done through algorithmic program that controls the system via using a reflection process (the **inference machine)** which is regarded as the responsible part of cognition in the system. The solutions provided by the system often need clarification and explanation, which are comparable to the demand for experts to explain their points of view, Therefore it is (the **explanation tool) that** allows the user to understand the logic and reasons behind the conclusions presented by the system. All of that carried out through the **interface** that enables the end user can to communicate with the system.

4. RESEARCH METHODOLOGY

The whole academy staff in 4 of public universities, and 3 of privet college universities was chosen as the research community. The location of the research includes the scientific departments that have relatives with the computer science field. The reason of choosing this sector as the sample is due to its importance and the high scientific (Knowledge) qualifications of its staff members, and the suitability of their specialties with the nature of the research variables. The research sample was represented by (70) Academic staff members from above sectors (distributed as 81 total questionnaire forms, where 6 were not returned and 5 were invalid due to the failure to fill all the required data) then a questionnaire was developed to verify the validity, as well as stability by using the method of (Spearman & Brown method) where the Reliability coefficient was (0.753). At the end; the data were examined statistically using the SPSS (Statistical Package for the Social Sciences).

5. FINDINGS AND DISCUSSION

1) THE DESCRIPTIVE STATISTICS RESULTS

Statistical methods were used to calculate the values of the research variables, which reflect the general trend of the sample response to the questionnaire, such as the mean, to estimate the mean response measurements for the sample items, and the standard deviation to get an idea of how homogeneous the values of this response, as follows.

The Information in the Engineering Knowledge System:

Table 1 shows that the departments that achieved measured mean of (3.397) which is very close to the standard deviation. Moreover these departments have achieved the highest mean in term of their Knowledge of the Hardware Components of the System (3.726) that means the scientific departments acquire good sum of information related to the Hardware components of the Knowledge Engineering System that exceeds their knowledge of the human components or the system operations.

Variables		Variables Code	Mean	standard deviation
Information	The physical components	V411	3.726	2.296
	The Human Components	V412	3.433	2.199
	Knowledge Engineering Operation	V413	3.033	2.219
		V41	3.397	2.029
Practices	The physical components	V421	2.70	1.98
	The Human Components	V422	3.28	1.93
	Knowledge Engineering Operation	V423	2.86	2.06
		V42	2.974	1.84
Experience in the physical components	Knowledge Base	V431	3.024	1.777
	Knowledge Acquisition Facility	V432	3.102	1.977
	Inferences Engine	V433	2.636	1.844
	Explanation Facility	V434	3.291	2.296
	User Interface	V435	3.605	1.953
	The Development engine	V436	2.584	1.690
	The Black board	V437	2.976	1.823
	Refining System Knowledge	V438	3.150	1.809
		V43	3.046	1.668
Experience in the Human Components	Knowledge Engineer	V441	3.297	2.012
	The Expert	V442	3.532	2.300
	The User	V443	3.527	2.030
		V44	3.452	1.909
Experience in Knowledge Engineering Operation	Knowledge Acquisition	V451	3.386	1.787
	Validation & Verification	V452	3.553	2.164
	Knowledge Representation	V453	3.136	1.804
	Inference Processes	V454	3.320	1.963
	Knowledge Explanation	V454	3.043	1.967
	Development of the system	V456	2.944	2.114
		V45	3.231	1.776
		V4	3.214	1.576

The Practice of **Engineering Knowledge System:**

As shown in table (1) that the mean of the knowledge engineering system was (2.947), standard deviation (1.840). and the highest average standard of the section practice of system users, with average value of (3.441) and standard deviation (2.349), this proves that Sections practices as system users exceeds other practices.

Experience of physical Components of the Knowledge Engineering System:
The response indicated that the mean in this area was (3.046), and a standard deviation of (1.668). Table 1 also shows means was in the field of "interface" (3.605), standard deviation (1.95), while the lowest was for the "machine development" with mean (2.584), and standard deviation (1.690). The responses varied indicators for other components between these two extremes figures.

Table 1 the descriptive statistics indicators results of the variables researched

The **Experience of Human Components of the Knowledge Engineering System:**
Table 1 shows the mean was (3.452), and standard deviation (1.909) in terms of expertise Possessed by sections in the field of knowledge of human components of the system. However the highest response is in the field of "expert" and "users" which are very close.

The **Experience of** Operations **of the Knowledge Engineering System**:
The results shown in Table 1 that the scientific departments mean were (3.230). The highest mean was of (3.553) for "process to confirm the reliability of the system" followed by "the process of concluding knowledge", while the "development" has achieved lower mean and lesser importance among System Operations.

2) RESEARCH HYPOTHESIS TESTING

- First **Hypothesis System Testing by Correlation:**

Pearson Technique was used to measure the strength of relation between each of the two research variables, to make use of the property of link-association (16) in the interpretation of the nature of the derived relationships between the variables researched. However, it is clear from Appendix 1 that "the dimensions of knowledge engineering system" have all made significance positive relationships, 5 of the relations were below significance ratio of (0.05). while (2.48) were below the level (0.01) out of total of the relations (253) and this is what supports the acceptance of first hypothesis.

- **Second Hypothesis System Testing by Factor Analysis**

In order to test the second hypothesis of the research, factor analysis has been used to reach to the simple factor architecture via "Hoteeling Principle Component analysis" for analyzing the relations array based on factors. On the other hand to determine the significance factors, "Kaiser Criterion" has been used, which is the most common method to identifying significance factors in factor analysis. Therefore, factor analysis has been used to achieve the following:

- Degree of inclusion of variables affecting the designed system.
- The sequence of these variables according to their importance.
- Assessing the degree of importance of studying a phenomenon according to these components.

Varimax with Kaiser Normalization
For the purpose of achieving estimates of the factor analysis, the research was dependence on initial Correlation Matrix, which shows clusters of linear Correlation relations between the researched variables, reflecting the high level of "Common Variance" within the overall contrast of variables system. A long with Table 2 shows the research variables communalities estimation, the highest value was for the variable "Inferenceing Operation" (0.912), and the least value was for the "Expert" (0.657). To find the Simple Factor-based Structure by using the components analysis of Hoteeling, extract factors whose Eigenvalues are less than one, axis have been rotated in order to get the structure, which led to the used on the Rotation of Axes Process for the purpose of diagnosing the Eigen Factors, according to the Varimax with Kaiser Normalization method.

Variables		Variables Code	Communalities
Information	The physical components	V411	0.784
	The Human Components	V412	0.759
	Knowledge Engineering Operation	V413	0.868
Practices	The physical components	V421	0.894
	The Human Components	V422	0.757
	Knowledge Engineering Operation	V423	0.840
Experience in the physical components	Knowledge Base	V431	0.803
	Knowledge Acquisition Facility	V432	0.790
	Inferences Engine	V433	0.662
	Explanation Facility	V434	0.745
	User Interface	V435	0.858
	The Development engine	V436	0.746
	The Black board	V437	0.805
	Refining System Knowledge	V438	0.742
Experience in the Human Components	Knowledge Engineer	V441	0.819
	The Expert	V442	0.657
	The User	V443	0.701
Experience in Knowledge Engineering Operation	Knowledge Acquisition	V451	0.881
	Validation & Verification	V452	0.762
	Knowledge Representation	V453	0.822
	Inference Processes	V454	0.912
	Knowledge Explanation	V454	0.815
	Development of the system	V456	0.769

Table 2 Variables Communalities Estimation

As Table 3 clarifies three factors, after seven rounds of succession, the percentage of (79.087%) was interpreted of the total variance value, where the highest Eigen value was (7.097), which formed (30.858%) of the total variance, and the lowest value was (5.464) as a percentage (23.759%) of the total variance ratio.

Table 3 also, demonstrates the loading of the first factor components by Rotation of Axes, the total number of loading components of this factor were (18) that represents (78.3%) of the total components that are candidates for analysis, while the Eigen value of this factor was (7.097) representing the percentage of Explained Variation percentage of (39.858%) of the total percentages (79.087%). The total number of high-impact loading components was (15), intermediate-impact (3). the saturated components of this factor($\pm$5) or more, that are shown in the Table 3, since they are highly related to the **importance of the experience in both the components of the system and its operations.**

The total number of components that are loading in the Second Factor by Rotation of Axes were (18) of loading components of the statistical significant and a percentage of (78.3) of the total components that candidate for the analysis, Table 3 further provides that Eigenvalues for this group was (5.628) and the percentage of Explained Variance was (24.47%) of the total number of the percentage (79.087%). Nine components have a significant impact on this factor, which mostly related to **the Practices of the system and the experience in it.**

Variables		Variables Code	communalities 1	2	3
Information	The physical components	V411	-	-	
	The Human Components	V412	0.327	-	
	Knowledge Engineering Operation	V413	-	0.353	
Practices	The physical components	V421	-	0.425	0.830
	The Human Components	V422	-	-	0.823
	Knowledge Engineering Operation	V423	0.332	0.517	0.680
Experience in the physical components	Knowledge Base	V431	0.495	0.584	0.466
	Knowledge Acquisition Facility	V432	0.608	0.408	0.505
	Inferences Engine	V433	0.603	0.416	0.392
	Explanation Facility	V434	0.666	0.515	-
	User Interface	V435	0.732	0.462	0.330
	The Development engine	V436	0.589	0.475	0.417
	The Black board	V437	0.775	-	0.360
	Refining System Knowledge	V438	0.790	-	-
Experience in the Human Components	Knowledge Engineer	V441	0.753	0.388	0.319
	The Expert	V442	0.751	0.300	-
	The User	V443	0.727	0.301	-
Experience in Knowledge Engineering Operation	Knowledge Acquisition	V451	0.724	0.675	-
	Validation & Verification	V452	0.542	0.675	-
	Knowledge Representation	V453	0.514	0.730	-
	Inference Processes	V454	0.502	0.801	-
	Knowledge Explanation	V454	0.330	0.823	-
	Development of the system	V456	-	0.772	-
		Eigen value	7.097	5.628	5.464
		Variances	30.858	24.47	23.759
		Cumulative	30.858	55.33	79.087

Table 3 Extracted Factors Matrix For By Rotation of Axes

The values less than ± 3 was blocked (have no value in the analysis and make the table easier to read)

It is also noted from Table 3 that the number of components that Loaded the third factor by using Rotation of Axes were (10) statistically significant components, and accounts for (43.48%) of the total components that are candidate for analysis, while the value of Eigenvalues for this factor (5.484)), and the percentage of the explained variance

(23.759%) of the total percentages (79.087%) and (4) components with a significant impact on this factor, which focused on **system Practices.**

So the second hypothesis can be accepted, which states the inclusion of the system with all dimensions researched and extract the main factors (high-impact, medium-impact variables).

6. CONCLUSION:

Through examining different aspects of the designed system in both its conceptual side and its testing; the following conclusions are made:

- The presence of different levels of knowledge (information, practice, experience) of sample members in regard to the system dimensions which reflects the possibility of its application.
- A relationship between the knowledge engineering system dimensions was found with significant correlation, and in an inverse way.
- The acceptance of all researched variables within the designed system dimensions.
- All variables form an important impact that is enough to join them to the rest of system dimensions.
- The capability of the designed system to extract knowledge from more than one expert in a specific field of specialization in a particular organization, and thus benefit from the differences in these views.
- The possibility of implementing more than one knowledge base in an organization for each knowledge area according to the knowledge activities of that organization.
- The development of information technology in the future could lead to more perfection of this system, thereby reducing the role of knowledge engineer, and the extract of knowledge from the expert directly.