In 3rd International Conference on the Practical Applications of Knowledge Management (PAKeM'2000), April 12-14, Manchester.

A Convergence of Knowledge Management and Data Mining: Towards 'Knowledge-Driven' Strategic Services

Syed Sibte Raza Abidi and Cheah Yu-N

School of Computer Sciences
Universiti Sains Malaysia
Penang, Malaysia.
Email: sraza@cs.usm.my

Abstract

Lately, there is a growing realisation amongst the healthcare community to leverage upon the vast quantities of healthcare data and enterprise-wide knowledge to transform it into value-added, 'decision-quality' knowledge, vis-à-vis *Knowledge-Driven Strategic Healthcare Services*, oriented towards healthcare management and planning. In this paper, we firstly present an integrated Knowledge Management Info-structure—a *Healthcare Enterprise Memory*—with the functionality to acquire, share and operationalise the various modalities of knowledge existent in a healthcare enterprise. Secondly, we focus on a specific component of our Healthcare Enterprise Memory—the *Knowledge-Driven Strategic Healthcare Services Info-structure*—that effectuates a confluence of Knowledge Discovery in Databases and Knowledge Management techniques. Functionally, the proposed Knowledge-Driven Strategic Healthcare Services Info-structure leverages on existing healthcare knowledge and data bases to derive decision-quality knowledge and then operationalises the acquired knowledge in terms of Strategic Healthcare Services. In conclusion, we argue that the proposed Knowledge-Driven Strategic Healthcare Services Info-structure is an attempt to rethink the possible sources of leverage to improve healthcare delivery, hereby providing a valuable management resource to healthcare policy makers.

1.0 Introduction

In today's knowledge-theoretic enterprise cultures, there is a premium in the operationalisation of enterprise-wide data—embodying knowledge about *what have we done* and *what have been the outcomes* of organisation-wide actions/processes—to derive *Empirical Knowledge*. The economic value of Empirical Knowledge is that it provides a window on the internal dynamics of an enterprise, allowing managers/policy-makers/analysts to infer 'inherent', yet invaluable, operative principles/values/know-how/strategies pertinent to an enterprise. In a Knowledge Management (KM) parlance, the origination of empirical knowledge pre-empts the definition of a novel *KM Services Portfolio*, that encompasses the migration of raw empirical data to empirical knowledge to (empirical) *Knowledge-Driven Strategic Decision-Support Services.* The vantage point of the aforementioned KM services portfolio—i.e. strategic decision-support services—is that it allows managers/policy-makers/analysts to device policies or make strategic decisions or predict future consequences by taking into account the actual outcomes/performance of the enterprise's current operative values—which may not necessarily be the same as the espoused operative values.

In a KM context, the transformation of raw empirical data to value-added strategic decision-support services, primes the realisation of interesting and innovative KM applications. Innovation draws from the fact that traditional (deduction- and analogy-based) knowledge acquisition techniques, catering for both tacit and explicit knowledge procurement, may not necessarily suffice to capture empirical knowledge from data. Henceforth, we will need to look into alternate ways to exploit induction-based knowledge acquisition techniques, akin to the practices of *Knowledge Discovery in Database (KDD)* and *Data Mining*, in tandem with KM techniques.

Knowledge Discovery in Databases (KDD) is defined as "non-trivial extraction of implicit, previously unknown, and potentially useful information from data" (Piatetsky-Shapiro and Frawley, 1991). KDD is a rapidly evolving discipline that uses tools and techniques from artificial intelligence—in particular, machine learning and neural networks—and statistics to discover 'empirical' knowledge from large databases, containing the findings and outcomes of organisation-oriented tasks and processes. Functionally speaking, KDD techniques inductively derive from organisation-wide data by suggesting novel perspectives to data, i.e. summarising the data, finding complex association rules between various data elements, identifying multiple combinations of data elements that influence a given list of outcomes, clustering similar data items to identifiable categories, classifying data elements into known categories, finding sequential patterns, predicting future values of certain data elements based on past values and so on.

In this paper, we propose a confluence of KDD and KM techniques to (a) procure the hitherto 'hidden' empirical knowledge from enterprise-wide data by 'mining' enterprise-wide databases and then to (b) generate a suite of data-derived, knowledge-driven strategic decision-support services. The underlying intentions of our approach are twofold: (1) to suggest novel ways and sources to enhance the knowledge assets of an enterprise, the incorporation of KDD within the KM envelope is a case in point; and (2) to design innovative info-structures that leverage upon acquired empirical knowledge to pro-actively draw out *on-demand* and *just-in-time* strategic insights/recommendations/predictions/analysis for top-level executives to abet strategic decision making and policy formulation. Instead of talking at a generic level, we have chosen the healthcare enterprise—due to its data-rich and information-centric nature—as a candidate enterprise that can benefit from the incorporation of KM techniques that we propose, moreover reference to the healthcare enterprise in our work provides the much needed context to elucidate our aforementioned approach and the resulting architectural blueprint of the systems being developed. In this regard, our conjecture is that a synergy between KDD and KM techniques can allow us to leverage upon healthcare data repositories to harvest strategic 'knowledge-driven' decision-making services for the healthcare enterprise. The basis of our conjecture is that we notice that the healthcare industry is saturated with both explicit and implicit knowledge—Table 1 shows a sample of the various types of knowledge that exist in a healthcare enterprise and can be captured using KM techniques; the different types of knowledge bases that can be created and administered within a healthcare enterprise; and the kind of knowledge-driven services that can be derived from the available knowledge bases. To conclude, we will like to point out that vis-à-vis our work, we have managed to demonstrate the confluence of KM and KDD disciplines in terms of the emergence of a new agenda about how healthcare management can be enhanced in the future. Here, concepts and technologies are impacting on each other.

Types of Knowledge	Types of Knowledge Bases	Types of Knowledge Services
• Medical Knowledge • Organisational Structure • Operational Workflow • Protocols/Guidelines • Medical Procedures • Business Rules • Medical Knowledge • Patient/Community • Staff Profile • Resource Inventory (Human, Equipment, Building, Etc.)	• Domain Knowledge • Protocol Knowledge • Workflow Knowledge • Policy Knowledge • Lessons Learnt Knowledge • Admission Knowledge • Delivery Knowledge • Performance Knowledge • Discussion Knowledge • Business Knowledge • Change Adaptation • Communication Knowledge • Enterprise Documentation	• Healthcare Entity Modelling • Transfer of Best Practices • Benchmarking • Audit Trails • Resource Scheduling • Product/Service Evaluation • Product/Service Accreditation • Policy Revisions • Workflow Revisions • Training Programs • Business Ventures • *Strategic Knowledge Services*

Table 1: A view of healthcare knowledge, knowledge bases and knowledge-driven services

The work reported here relates to our present involvement in the Malaysian Tele-Medicine initiative (Government of Malaysia, 1997; Abidi *et al.*, 1998) under the auspices of the proclaimed *Multimedia Super Corridor (MSC)* project. For all intents and purposes, the Malaysian Tele-Medicine initiative not only envisages but deem imperative the need for a knowledge-centric healthcare environment that derives 'capital', 'intelligence' and 'efficiency' from the exploitation of rich (healthcare) knowledge resources—vis-à-vis the generation of knowledge-driven decision-making services—for healthcare management and strategic planning (Abidi, 1999; Abidi and Yusoff, 1998). In this regard, our work entails the development of such an integrated *Knowledge Management Environment* that, amongst other functions, also supports a suite of *Knowledge-Driven Strategic Healthcare Services (KSHS),* derived from nation-wide healthcare knowledge and data. Here, in this paper:

(1) We will present the architectural blueprint of integrated Knowledge Management Environment—a *Healthcare Enterprise Memory (HEM)*—that purports the functionality to acquire, share and operationalise the various modalities of knowledge existent in a healthcare enterprise (Cheah and Abidi, 1999a & 1999b). The proposed HEM projects a holistic knowledge outlook, incorporating (a) a wide variety of knowledge bases containing knowledge derived from both typical AI-based knowledge acquisition and Data Mining techniques; and (b) a variety of knowledge-driven strategic services.

(2) We will discuss the Knowledge Discovery, in particular Data Mining, component of HEM—the *Knowledge-Driven Strategic Healthcare Services Info-structure* that leverages on existing healthcare knowledge/data bases to (a) derive decision-quality knowledge from healthcare data and (b) deliver a suite of KSHS. Note that in this paper we will only be presenting the architectural blueprint of the KSHS Info-structure, with minimum coverage of its functionality.

(3) We will conclude that the proposed synergy between KM and KDD is a step forward in the right direction (Abecker *et al.*, 1998; Dorfman, 1997), as we will highlight that the proposed KSHS Info-structure—the solution to the procurement of the hidden knowledge contained in the seemingly mundane health data—can serve as a valuable decision-support resource to healthcare policy makers by providing them an intimate understanding of the healthcare enterprise.

2.0 Knowledge-Driven Strategic Healthcare Services: An Overview

The health care industry continues its rapid and remarkable transition from a fragmented industry to an extended enterprise, powered by health-related knowledge and information resources and driven by intense pressures to deliver cost-effective quality healthcare services. The effective delivery of healthcare services hinges on the ability to deliver appropriate and proactive value-added services to different client segments on a timely basis—indeed, the impact of services will be lost if the healthcare industry does not 'deliver the right service to the right client'—irrespective of the access enablers, distribution channels and technology employed. In general, healthcare services need to be systematically determined based on needs; packaged according to usage patterns, demographics and behavioural psychographics; and delivered in a ubiquitous, proactive and continuous manner. These mutually inter-related constraints are hard to formulate, let alone satisfy, using conventional strategic planning techniques. Notwithstanding the fact that the healthcare industry is quite complex with wide variations in practices, nevertheless there do exist many formal processes that can be analysed and measured in terms of definitive process models. In fact, the healthcare enterprises collect massive amounts of data that derive from such formal processes. Thus, there is a strong case for operationalising the seemingly placid healthcare knowledge and patient-level data into value-added, 'decision-quality' knowledge, vis-à-vis, (empirical) knowledge-driven decision-support services oriented towards healthcare management and planning. The rationale is that by understanding what worked—or did not—healthcare practitioners can identify areas for improvement and/or capitalise on past successful methods.

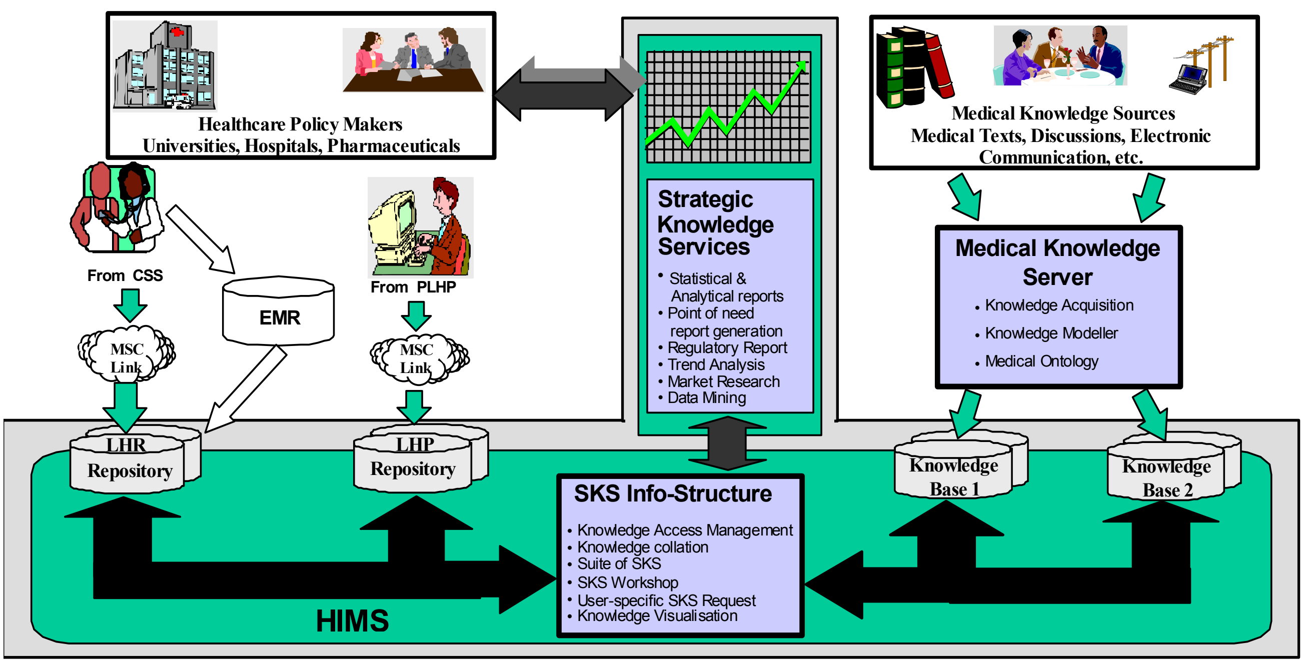

Figure 1: A pictorial overview of the KSHS environment.

Knowledge-Driven Strategic Healthcare Services (KSHS) can best be defined as a suite of knowledge/data-driven strategic services to facilitate the derivation or extraction, of 'decision-quality' information/knowledge/wisdom from (a) healthcare data and (b) the health enterprise's knowledge bases, leading to improved delivery of quality healthcare services (see Figure 1 for a typical KSHS environment). Methodologically speaking, KSHS cater for the migration of *data* (the what), through a sequence of sophisticated operations to *information* (the trend or behaviour) to *knowledge* (the why) to the ultimate level of *wisdom* (the necessary actions to be taken). The innovative meta-level of wisdom provides an additional dimension to KSHS where, the data collected, the information determined, the knowledge acquired can be used towards the subtle shaping of the healthcare enterprise so as to acquire the highest degree of healthcare quality by either fine-tuning the existing enterprise-wide practices or by defining new action plans for policy making and health planning (Abidi, 1999; Abidi and Yusoff, 1998).

Typical KSHS may include: data mining based services for helping health-care providers cut costs and improve care by indicating which treatments statistically have been most effective; benchmarking for knowing how various physicians and hospitals compare with their peers; outcomes measurement to identify people statistically at risk for certain ailments so that they can be treated before the condition escalates into something expensive and potentially fatal; trend analysis of diseases/epidemics (Abidi and Goh, 1998a and 1998b), treatment patterns, hospital admissions, drug patterns and so on; what-if scenario analysis; comparing medical practices with medical business rules; market research; point-of-need generation of statistics; publication of regulatory and best-practices reports; feedback routing to R&D institutions (e.g. drug effectiveness on outcomes of treatment); analysis for healthcare financing; data analysis for health surveillance; examining resource allocations and so on (Government of Malaysia, 1997). Table 2 gives an abridged list of possible KSHS.

Vis-à-vis KSHS, such a huge collection of critical health data is envisaged to ensure that the healthcare enterprise delivers the right service at the right time to the right user group. KSHS will eventually have a significant impact towards better health planning; segmenting and targeting medical products; evaluating the effectiveness of health related programs; focusing the delivery of medical services so as to be more proactive and effective; and so on. Principal beneficiaries of KSHS are envisaged to be the Ministry of Health (MoH), pharmaceuticals, medical service organisations, private health providers, universities, and community health organisations.

Albeit the definition of the intended functionality of KSHS, it is yet premature to assess the overall impact of KSHS. Nevertheless, it can be safely argued that as is true with any emerging trend, the ultimate factor that determines its adoption is its ability to drive economic value—value that will be measured by cost efficiencies, knowledge growth, and the ability to help the enterprise effectively leverage their knowledge assets. The same is applicable for the proposed trend of KSHS. For all intents and purposes, only the complete implementation of the KSHS info-structure, followed by the pro-active generation of strategic services will determine the ultimate efficacy of our approach! Yet, it is gratifying to note that the conception of KSHS has offered an intriguing dimension to the role of KM in healthcare.

Main KSHSs	Specific KSHSs
1. Analysing trends in hospital admission	1.1 Spectrum of disease 1.2 Seasonality in disease pattern 1.3 Interventive measures to be instituted
2. Analysing treatment pattern	2.1 Comparison between hospitals and within hospitals 2.2 Auditing against acceptable localised treatment consensus 2.2.1 criteria for admission 2.2.2 investigations 2.2.3 therapeutic intervention
3. Analysing outcomes of treatment	3.1 Gauge standard and quality of care 3.2 Enhancing Health System Research 3.3 Highlight aberrations in treatment outcomes 3.4 Develop norms/ standard of outcome measures acceptable to local realities
4. Analysing cost-effectiveness of health care	4.1 Audit expenditure and income 4.2 Highlight areas to focus 4.3 Allows manpower planning 4.4 Allows planning for infrastructure development
5. Planning out-of hospital (ambulatory) care	5.1 Assessing client (patient's and relative's satisfaction 5.2 Allows plan of care after discharge to be made 5.3 Monitor treatment after discharge and out patient's management
6. Forecasting 'new disease' and strategising appropriate preventive measures	6.1 Warm health planners of impending epidemics 6.2 Allows appropriate preventive strategies to be recommended at community level 6.3 Better public education strategies
7. Forecasting complications of treatment	7.1 Hospital acquired infection 7.2 Drug resistance pattern 7.3 Iatrogenic diseases

Table 2: A synopsis of the proposed data-driven strategic healthcare services

3.0 The Healthcare Enterprise Memory: A Four-Layer Environment

The practice of healthcare directly and crucially impacts the well-being of the community. For that matter, the healthcare community demands that all new healthcare services/procedures be firmly grounded into well-established medical knowledge and formalisms. Since KSHS, though being novel in nature and context, are envisaged to impact the content and delivery of healthcare services, we argue that KSHS should not be ad hoc in nature. Rather, KSHS should be based on a well-defined KM environment that supports the correct, transparent and systematic transformation of data to knowledge to strategic services. More so, the KM environment should address issues ranging from knowledge acquisition to knowledge sharing to knowing operationalisation. We believe that adherence to a pre-specified KM environment will ensure that the KSHS are grounded in well-defined healthcare (knowledge) rules, and henceforth the eventual outcome of the KSHS will manifest as tangible, opportune, realistic and most importantly functional suggestions that will keep the healthcare enterprise synchronised with the ever-changing health requirements.

To establish the validity and efficacy of KSHS we present an integrated KM environment, i.e. a *Healthcare Enterprise Memory (HEM)*—akin to an *Organisational Memory*—that provides the functionality to acquire, share and reuse the various modalities of knowledge—tacit and explicit knowledge of healthcare practitioners, healthcare related documents, healthcare data, healthcare processes and workflows, experiences, lessons learnt and so on—within a healthcare environment The proposed HEM supports a number of knowledge-oriented services, such as automatic dissemination of knowledge, reuse of knowledge and experience, support of intelligent knowledge management services, timely provision of knowledge and experience, repository of healthcare knowledge of healthcare experts (Euzenat, 1996), transforming information to action, connecting and converting knowledge and above all ***healthcare modelling*** (Cheah and Abidi, 1999a and 1999b). Functionally, the main objective of HEM is to enhance the enterprise's competitiveness and efficacy by way of managing its knowledge and generating a suite of knowledge-driven services (Abecker *et al.*, 1998), i.e. the earlier discussed KSHS.

Put simply, the implementation of the so-called HEM is a non-trivial and perpetual activity that involves the active participation of a variety of players, more importantly a team of KM and healthcare practitioners. HEM developers are actively involved in the process of identifying and capturing healthcare knowledge from internal and external sources and storing it in knowledge-bases (see Figure 2). The various modalities of healthcare knowledge identified by us are: (a) formal knowledge (from texts and documents), (b) informal knowledge (from experience and lessons-learned) and (c) empirical knowledge (from data warehouses and databases). Among suggestions in realising HEMs are *lessons-learned archives*, *distributed case bases*, *expert systems and formal knowledge structures*, and *formal representations of argumentation*. In summary, the healthcare knowledge bases are to be populated by the abstraction of internal and external healthcare information/knowledge based on certain content identification criteria. The synthesis process transforms the abstracted knowledge into a formal representation scheme that renders it operable by KM systems. Finally, the emergent knowledge bases undergo a process of review or update to ensure the validity and consistency of the extracted knowledge (see Figure 2).

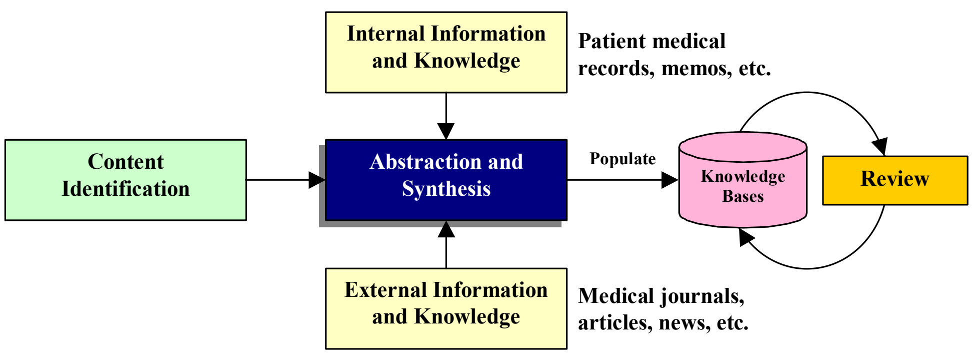

Figure 2: A Knowledge Base Creation Environment

For our purposes, the proposed HEM comprises four functionally distinct layers (see Figure 3):

1. *Object Layer*: Consists of various healthcare information and knowledge sources such as data, document, knowledge and scenario bases. The sources may have both formal (machine-readable) or informal (human-readable) representations.
2. *Knowledge Description Layer*: Enables uniform and intelligent access to object-level resources. The main purpose of this layer is to facilitate accurate specification and selection of relevant healthcare knowledge pertaining to the context. For this purpose, ontologies need to reside at this layer to maintain a standard vocabulary to describe concepts and relationships between

entities that attempt to share knowledge (Chandrasekaran and Josephson, 1999; Gruber, 1993a and 1993b).

3. *Application Layer*: Models and executes processes and tasks. The HEM's services can be realised in different ways, ranging from dedicated programs (which perform a well-defined task) to flexible query interfaces. These include medical protocol models and healthcare work processes management systems.
4. *Services Layer*: Provides specialised services to healthcare professionals or the public through the use of various applications. For instance, the KSHS portfolio akin to the group data services proposed by Abidi (Abidi, 1999; Abidi and Yusoff, 1998). Examples of these services include forecasting, knowledge and trend analysis, and best practice reporting services.

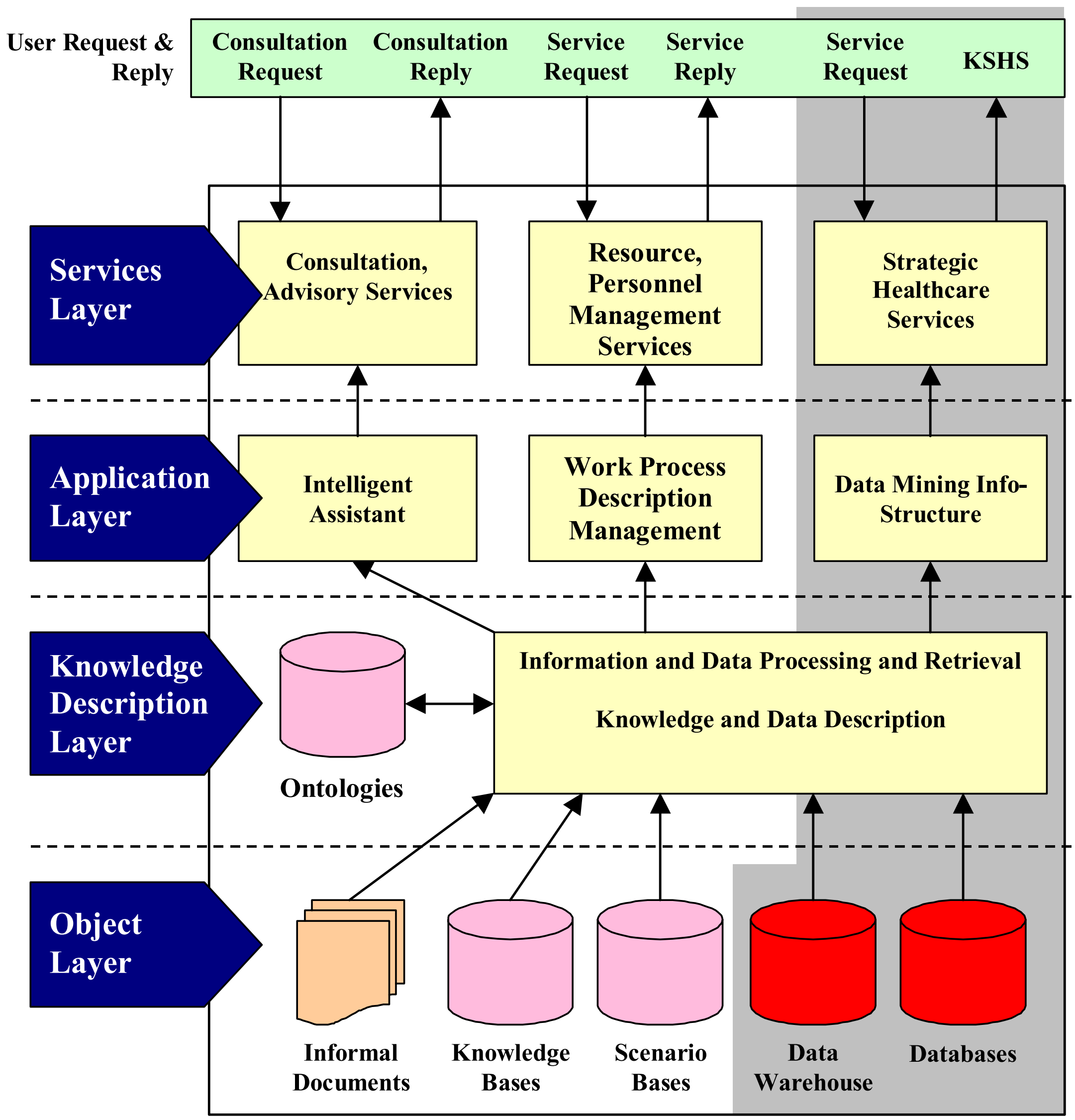

Figure 3: The Healthcare Enterprise Memory Model

The Object Layer holds an important role in the HEM as the entire knowledge, information, data resource in the healthcare enterprise are stored at this layer. Databases and Document Bases store the bulk of the healthcare enterprise's data and electronic documents while Knowledge Bases store decision support rules. However, the key component of the Object Layer is the Scenario Base. It stores the enterprise's tacit knowledge. A complete scheme of a HEM and their peripheral technologies and services is illustrated in Figure 4, whereby we show a mapping of the four-layer of the HEM model to KM processes, to KM technologies and to various AI technologies.

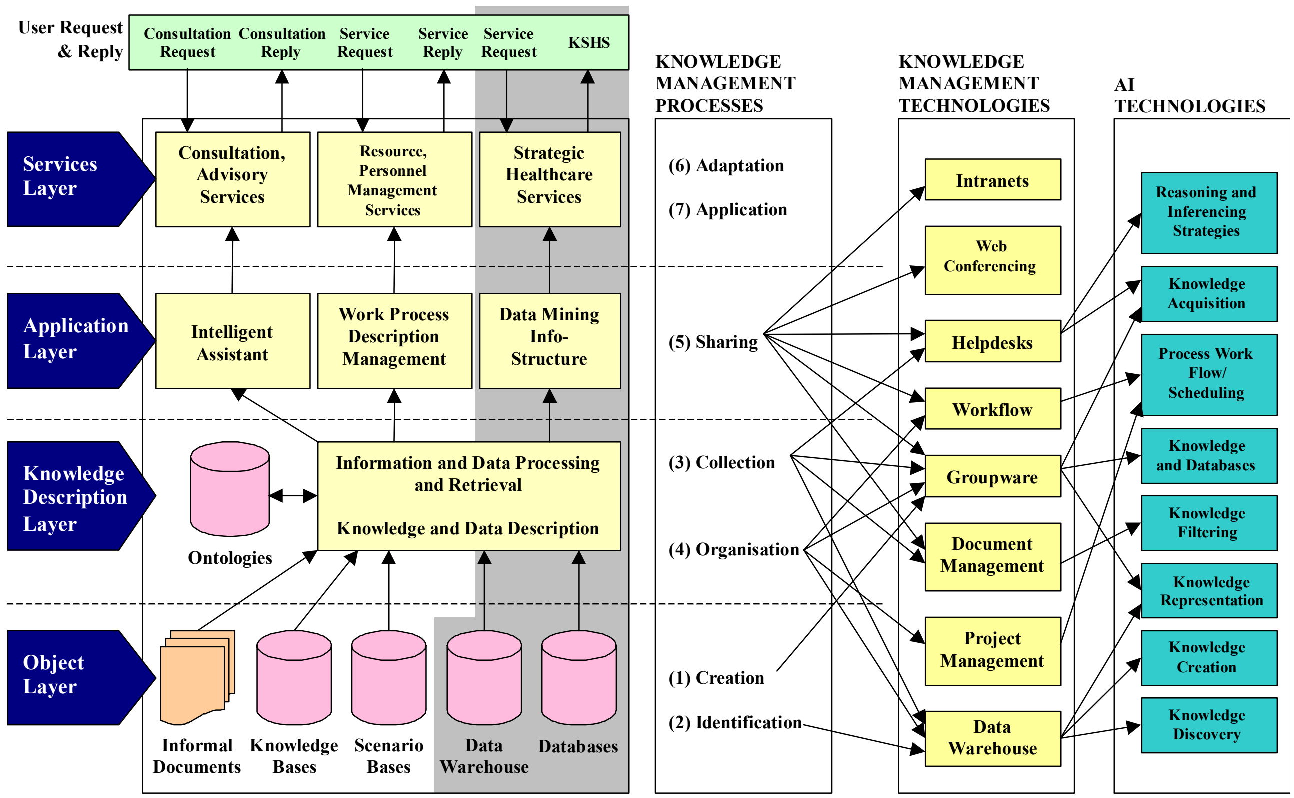

Figure 4. The Healthcare Enterprise Memory, showing the mapping of the four-layers to KM Processes to KM Technologies and finally to AI Technologies

In conclusion, it is our belief that the proposed knowledge management model—i.e. HEM—will launch the evolution of a healthcare enterprise through a hierarchical definition of KM tasks implemented in the HEM application portfolio. One such task is KSHS. As part of this evolution, it is anticipated that the healthcare enterprise will move from healthcare practice rules that are primarily stated in texts (within policy and standards documents, practising manuals and requirements specifications) to a knowledge-driven environment, capable of directly impacting the behaviour of the healthcare enterprise.

In this paper, we primarily focus on the KSHS info-structure, shown as the shaded area in Figure 3, implemented within the HEM. The forthcoming discussion will specifically focus on the architecture of the KSHS's info-structure, its functionality and inherent features. Discussions on the other components of HEM can be found in related papers (Cheah and Abidi, 1999a; Abidi, 1999)

4.0 The KSHS Info-Structure: Architectural Details

Our proposed KSHS info-structure derives from the KM methodology discussed earlier, thereby manifesting a suite of modules that facilitate the transformation of data to knowledge to services. Conceptually, the KSHS info-structure has a multi-tier architecture that effectuates a confluence of KM and data mining techniques. Figure 5 is an overview of the 4-tier architecture of the KSHS info-structure.

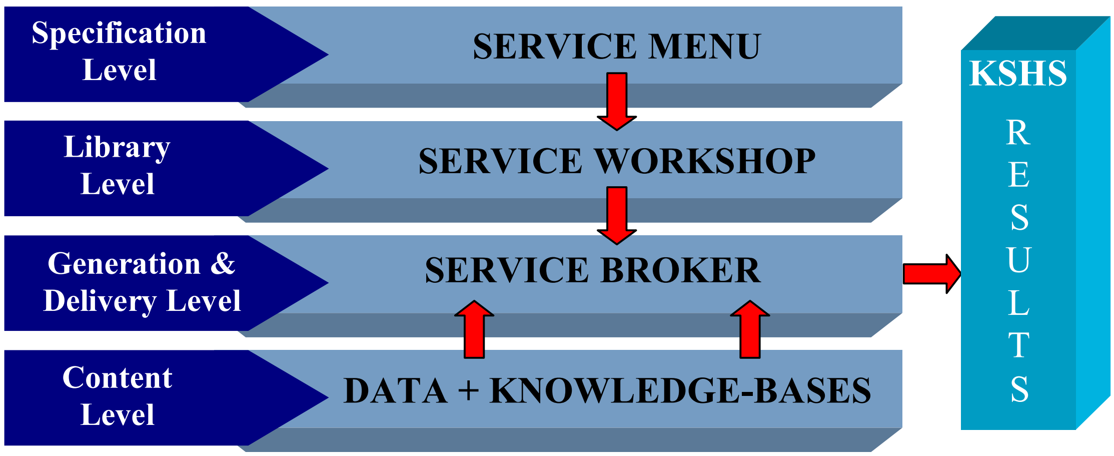

Figure 5: An overview of the multi-tier architecture of the KSHS info-structure

In procedural terms, the scheme for the provision KSHS is quite simple: (1) specification of the user's desired strategic service(s); (2) procurement of the necessary knowledge and methods (in terms of KSHS modules) relevant to the KSHS's specification; (3) generating the specified strategic service by the systematic manipulation of the KSHS modules in concert with data and knowledge from the data and knowledge repositories; and finally (4) the delivery of the KSHS results to users.

The multi-tier architecture of the KSHS info-structure (as shown above in Figure 5 and further elaborated in Figure 6 shown below) is reflective of the above-mentioned sequence of operations. The KSHS info-structure can be visualised as having four distinct levels, characterised by the roles/aspects pertaining to the overall processes underlying the generation and delivery of KSHS. The top level—the *Specification Level*—is responsible for the specification of a particular KSHS by users. The second level—the *Library Level*—stores knowledge pertinent to the provision of a variety of KSHS; each designated KSHS is stored as a specific *KSHS module* (encapsulating the processing methods and knowledge). Next, the third level—the *Generation & Delivery Level*—is the main workplace (i.e. the info-structure's engine) where the strategic services are generated by the execution of the KSHS modules (as per their processing script) in tandem with pertinent data and knowledge drawn from a variety of databases and knowledge-bases. The results/findings/conclusions/recommendations of the concluded KSHS are then delivered to the users. Finally, the *Content Layer* houses the enterprise-wide data and knowledge stored in one or more databases and knowledge-bases.

We now present the functional details of the various components of our KSHS info-structure. For the purpose of this paper, we will focus more on the 'data-specific' aspects of the info-structure (in particular issues pertaining to the access of data from data repositories). For all practical purposes, it may be noted that the implementation and delivery of KSHS involves an active interplay between both the data-specific and knowledge-specific components of the info-structure. Figure 6 gives the architecture of the proposed KSHS info-structure.

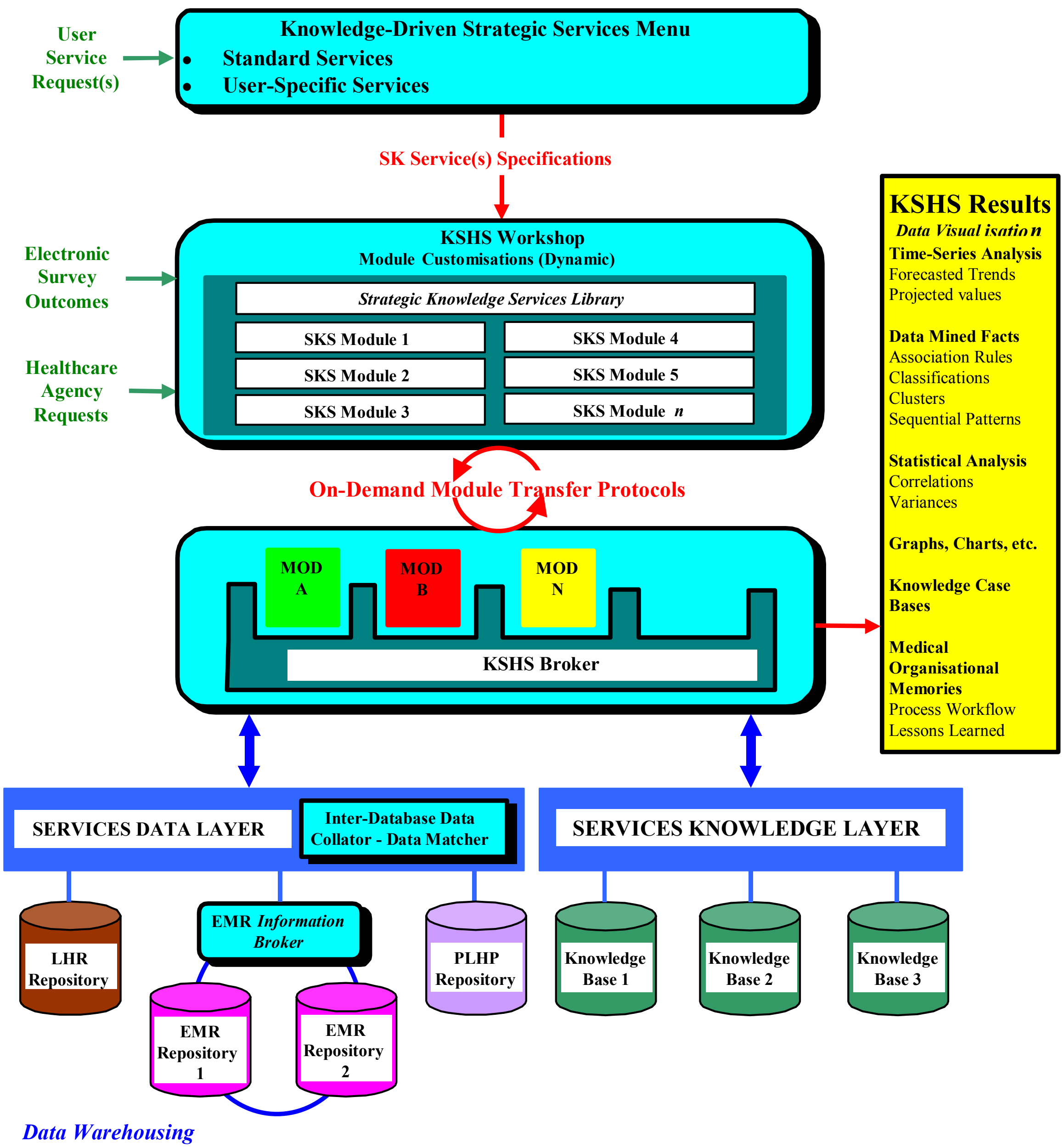

Figure 6: The architecture of the KSHS info-structure

Architecturally, our proposed KSHS info-structure manifests a suite of 'KSHS-modules', where each module is module provides a particular KSHS service by way of employing a designated data mining algorithm/technique and data visualisation technique. In principle, a KSHS-module facilitates the transformation of data to knowledge to wisdom. In this regard, the KSHS info-structure is an amalgamation of data mining algorithms pertaining to association rules, classification, clustering, sequential patterns, trend analysis, predictive modelling and so on. Again, a variety of computer technologies/techniques are exploited to meet the above mentioned tasks, including neural networks, machine learning, symbolic artificial intelligence, time-series analysis, statistics, etc.

4.1. Knowledge-Driven Strategic Healthcare Services Menu (KSHS-Menu)

KSHS-Menu is a front-end application that serves as the GUI for the specification of all types of KSHS. With built-in intelligence, KSHS-Menu will support script based specification that will guide the user through the entire specification process by self-prompting and recording their responses. We propose two modes of KSHS specifications:

(1) The user may choose a specific KSHS from a pre-defined services menu.
(2) The user may 'design' a specific KSHS by the dynamic synthesis of multiple standard KSHS. This option will demand from the user to (a) define a new service; (b) choose the pertinent attributes from multiple services; and (c) strategically align them together to form a new KSHS. The *Module Customisation Workbench* (discussed later) will next be used to develop the required KSHS-module as per user specification. Various considerations arise with this option and we intend to exploit 'intelligent' technologies for the delivery of this innovative option.

4.2. Knowledge-Driven Strategic Healthcare Services Workshop (KSHS-Workshop)

For all practical purposes, the suite of KSHS need to address the ever-changing needs of the healthcare enterprise. In view of this possibility, it is not viable to have just a limited set of KSHS, rather what is needed is an elastic framework to 'dynamically' tailor existing KSHS modules to develop new KSHS modules catering for customised (even temporary) services.

To meet the above objectives, the KSHS info-structure incorporates the *KSHS-Workshop* which is an application environment that supports (1) the storage of all KSHS modules in a *Knowledge-Driven Strategic Healthcare Services Library (KSHS-Library)*; and (2) the generation of demand-specific, new KSHS modules by the *Module Customisation Workbench*.

KSHS-Library, as the name implies, is an infrastructure housing a set of KSHS modules. Our proposal entails that each individual services be implemented as an individual self-contained KSHS-module. For instance, three KSHS such as (1) analysing the trends in hospital admission (2) analysing the trends in treatment patterns and (3) analysing the trends outcomes of treatment, will be implemented as three individual KSHS-modules. The efficacy of this approach is justified by two facts: (a) each KSHS is therefore identified as a unique, non-overlapping and self-contained entity; (b) the 'library' based infrastructure supports the addition of new KSHS without disturbing the existing composition of the library and the internal dynamics of the available KSHS-modules.

In technical terms, each KSHS module will be configured along the following principles:

- The scope of the service.
- Input data elements needed.
- Knowledge required to manipulate the data. The knowledge may either be directly 'hard-coded' within the processing 'script' of the module or else may be retrieved from the available KBs.
- Specification of the possible relationships between the data elements.
- The data processing and analysis methods. For instance, time-series analysis, data mining, etc.
- Specification of the output (service's outcome).
- Service's outcome visualisation methods.

Module Customisation Workbench (MCW) is an application that provides the functionality to (1) take user requests for specialised KSHS and then (2) 'dynamically' develop new KSHS modules based on existing ones. Three possibilities for the realisation of new KSHS modules are:

1. Mixing existing KSHS-modules in some principled manner to realise a 'hybrid' KSHS-module (shown in Figure 7). In this case, either entire modules are synthesised or else specific elements of each module are synthesised (we believe the latter approach is more practical).
2. The tuning of existing KSHS with regards to community feedback received in terms of electronic surveys. Here, we are proposing a mechanism that allows feedback from the healthcare agencies and the community to be included in the decision-making and strategic planning cycle.
3. Healthcare agencies need to deal with changing healthcare demands. To tackle atypical situations, our proposal provides a mechanism for healthcare agencies to submit specific requests for atypical processing and analysis of the existing data. Here, the MCW can systematically synthesise existing KSHS-modules to yield the desired services (shown in Figure 8).

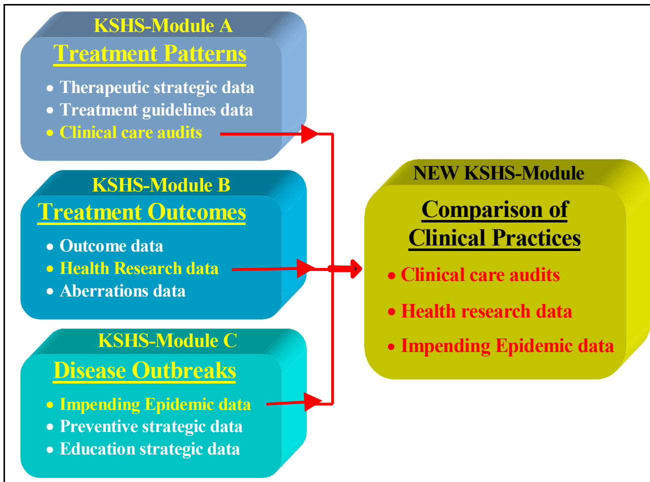

Figure 7: The generation of a new KSHS-module by combining the data elements from multiple existing KSHS-modules. Data/knowledge representing *Clinical care audits* (from KSHS-module A), *Health research data* (from KSHS-module B) and *Impending epidemic data* (from KSHS-module C).

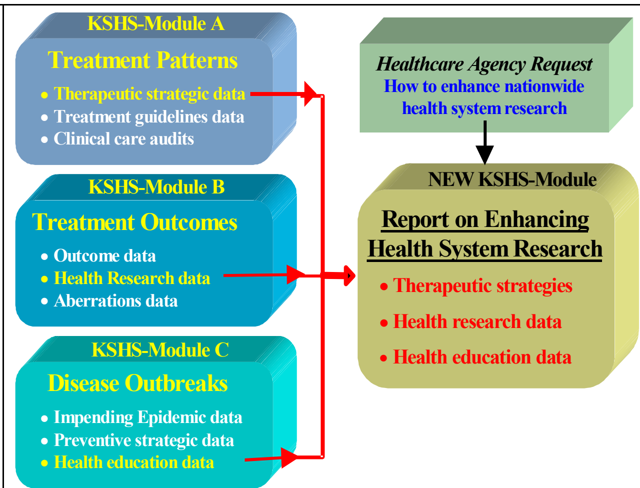

Figure 8: The generation of a new KSHS-module in response to a request from a healthcare agency. Data elements from multiple existing DSHS-modules are selected to realise a new DSHS-module.

4.3. Knowledge-Driven Strategic Healthcare Services Broker (KSHS-Broker)

We mentioned earlier that our proposal accounts for dynamic (both macro and micro) customisations of existing KSHS to meet atypical user requirements. To meet this end, the KSHS-Broker module provides the functionality to dynamically 'mix and match' multiple existing KSHS-modules to deliver innovative, need-of-the-hour services customised according to atypical user requirements.

We regard the KSHS-Broker as the engine of the KSHS info-structure, as it is responsible for the generation and delivery of the specified KSHS. In procedural terms, the KSHS-Broker builds upon the KSHS's specification, the necessary healthcare data and the pertinent knowledge to provide the stipulated strategic knowledge services. Architecturally, the KSHS-Broker can be envisaged as an infrastructure comprising a number of independent 'slots', where each slot can be filled by a KSHS-module. The eventual KSHS generated by the KSHS-Broker derives from the systematic amalgamation of the multiple KSHS-modules within the different slots. Hence, the slot based architecture of the KSHS-Broker predicates the possibility of dealing with an assortment of KSHS-modules, all at the same time, to yield a more elaborate and sophisticated output. The slot-filler analogy is highly relevant here to elucidate the functionality of the KSHS-Broker.

Indeed, the dynamic mixing of multiple KSHS-modules may lead to functional complications. But, the strength of our proposal lies in the fact that we have addressed this matter at the infrastructure level (by incorporating necessary KSHS synthesis mechanisms) and not at the KSHS-module level. Typical issues addressed at the design stage of the KSHS-Broker are:

- scheduling protocols for each module to access the services data and services knowledge layers;
- the 'hooks' to assimilate the modules in the slots;
- the sequence in which the modules will be visited by the processing engine;
- the compilation of the analysis by each module to generate a global report.

It may be noted here that, there exists an *on-demand module transfer* protocol between the KSHS-Workshop and the KSHS-Broker. A cache system is the key to this protocol, whereby the module(s) needed by the KSHS-Broker are placed in the cache and are delivered to it as and when needed, i.e. an *on demand transfer protocol*. KSHS-modules which are no longer needed by the KSHS-Broker are swapped out to the KSHS-Workshop. This approach ensures that the KSHS-Broker will not be over-loaded with modules that are not currently in use, thus eliminating concerns about time delays and memory shortages.

4.4. Services-Data and Services-Knowledge Layers

The services-data and services-knowledge layers are the backbone interface between the data/knowledge warehouse entities and the other processing components of the KSHS info-structure. In the context of data, the data-services layer is responsible for accepting the data request specification from the KSHS-Broker and subsequently delivers the required data from the data warehouse(s). Likewise, when the KSHS-Broker needs to draw upon knowledge stored in the knowledge bases, the services-knowledge layer provides the necessary access to the knowledge repositories. Important responsibilities of the services-data layer include (a) maintenance of data integrity, (b) data normalisation (if needed), (c) data cleansing (if needed), (d) extrapolating/interpolating to fill in for missing data values and (e) maintenance of data dependencies, relationships and data dictionaries. Here we discuss in detail the activities of the services-data layer. Other than the various data repositories, the services-data layer works in concert with two other data-delivery applications - *Information Broker* and the *Inter-database Data Collator and Data Matcher*.

Information Broker (IB) (Abidi, 1997) is used to collect and collate relevant information from the multiple data repositories. It supports a customisable environment for complex database navigation, data extraction and inter-database data collation and export. IB provides a customisable data access environment that spans across multiple databases (where each database could be of a different format and product platform). Equipped with an open-ended interface, IB can dynamically connect with any database product (Oracle, MS Access, FoxPro, etc.) and any database design. In real-time, users can alternate between multiple databases and perform database navigation and data collection. In summary, the IB dynamically allows users to connect to as many databases as deemed necessary; users can then perform customised transversals through databases to collect relevant data items;

finally, the data extracted from multiple databases is collated and made available for export to the services-data layer.

Inter-Database Data Collator And Data Matcher (IDDCDM) is an optional data-specific application that can be used to determine whether two data items in the same/different database(s) belongs to the same entity, i.e. to identify duplicate records (though they may be explicitly similar). The principal role of IDDCDM is to take a specification of data that need to be collected from multiple data repositories and then 'sift' through multiple databases to collect the required data item (even though multiple instances of the specified data may not be similarly explicated). In the next step, it determines record duplication, based on pattern matching and semantic analysis. In a healthcare context, such functionality is highly relevant as it allows us to deal with healthcare data (say patient information) originating from multiple sources, even though if the data is not explicitly defined.

4.5. Knowledge-Driven Strategic Healthcare Services Visualiser (KSHS-Visualiser) - Results Visualisation Module

The efficacy of the KSHS info-structure hinges on its ability to explicitly illustrate the data/knowledge analysis results. To meet these objectives, the KSHS-Visualiser module implements a combination of result visualisation techniques with an emphasis on graphic illustrations. Since the KSHS info-structure provides a variety of data/knowledge analysis techniques (such as data mining, time-series analysis, benchmarking, etc.), where each technique has its own peculiar reporting style, the KSHS-Visualiser module offers reasonable flexibility to the effect that users can determine how they want to view the results; users can customise report formats and screen displays during a simple (data analysis) result specification session.

KSHS-Visualiser supports the following reporting formats—graphs, tables, maps, abstract maps and visualisation hypercubes. We give below a brief description of each visualisation technique.

<u>Graphs</u>: Graphs will be used when the results portray a changing relationship between the properties involved.
<u>Tables:</u> For (tabular-numeric) reports that intend to determine the exact value for a particular property.
<u>Maps:</u> Our solution provides mapping - a popular technique for visually depicting information that explicitly portrays relationships (See Figure 9). Features of maps that make them especially good data visualisation aids are:

1. The dimensionality of the location information about the data is preserved in the dimensionality of the location information on the computer screen.
2. The map's quantitative properties - such as the relative distance between points can be processed with minimal effort and correspond to quantitative properties of the data.
3. Colours and shades can be used to visualise the content of a property.

<u>Hypercubes</u>: For higher-dimensional visualisation, a greater percentage of results needs to be presented in a manner that facilitates the comparison of complex patterns. To meet such demands we explicate high-dimension data using hypercubes - a grid-like structure (see Figure 10).
<u>Data Mining Results:</u> Typically, data mining results are classified into four classes:

1. Associations: Given a collection of items and a set of records that contain some number of items, an association result illustrates the affinities that may exist among the collection of items.

2. Sequential Patterns: Results of a sequential pattern function depict a set of frequently occurring patterns detected in related records.
3. Classifiers: The classification results will produce a description of the typical characteristics of all records that are the member of a class.
4. Clustering: The clustering results will depict a collection of records that could be clustered together on the basis of the similarities of the attributes.

Time-Series Forecasting and Trend Analysis: Results will best be represented as multi-dimensional graphs supplemented by a concise, yet information-rich, report which basically illustrates inherent trends in the data.

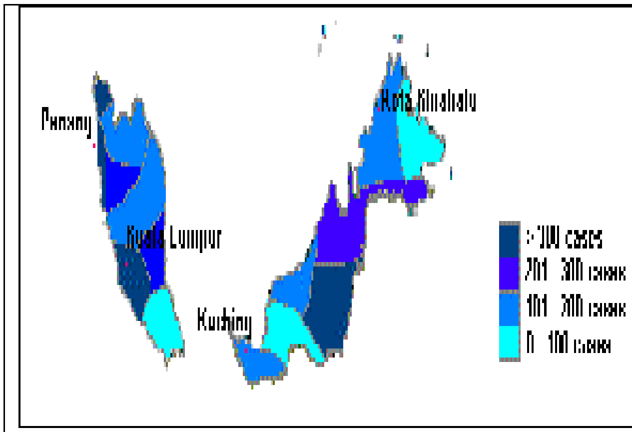

Figure 9: Using Maps to represent KSHS

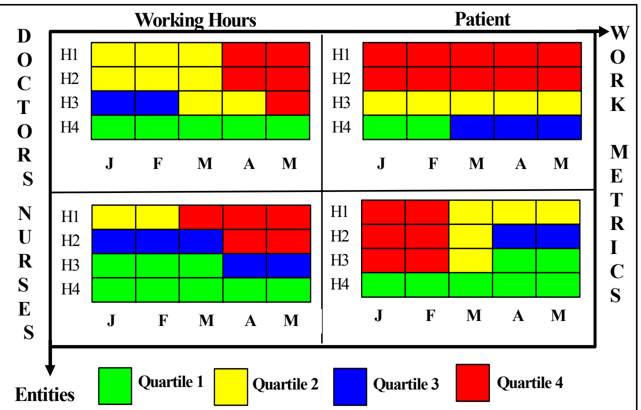

Figure 10: Hypercubes to represent KSHS

4.6. KSHS Delivery

Two intuitive, easy-to-use, visual applications are used to deliver the generated KSHS:

On the Internet: Standard information will be available on our web server and registered customers may choose to utilise the available KSHS. Reports may be delivered via the Internet which would then allow the dynamic generation of HTML documents. The delivery of on-demand health information will also be possible through our Internet service offerings; and for more advanced Internet customers requiring sophisticated analysis on the information from KSHS, we offer web-enabled ROLAP functions for trend analysis, statistical analysis and forecasting.

Direct Links to DSHS info-structure: This mode of delivery will be favoured by major corporations that require regular healthcare information to support their business operation, for example Insurance and Pharmaceutical companies, hospitals and MoH. The proposal is for a direct telecommunications link to be established to enable these customers to access standard information and also customise information that has been requested. As on the Internet, direct link customers will also have the capability to perform complex analytical functions using ROLAP applications. The direct link provides a targeted and proactive delivery channel

5.0 Features of the KSHS Info-Structure

In its most generic manifestation, KSHS is the scientific analysis of health data to derive decision-support information. Functionally, the proposed KSHS info-structure not only meets the particular demands for KSHS, but extends further to add value by way of supporting a number of attractive and innovative features, such as:

Data Collation: Comprehensive data analysis by collating data from multiple data repositories—where each data repository may contain logically different health-related information—to form a virtual, seamless and continuous data source.
Suite of Strategic Knowledge Services: A wide range of KSHS that address the demands of multiple clients—healthcare agencies, insurance and the pharmaceutical industry. The KSHS are categorised into four prominent classes (whereas within each class we have numerous sub-KSHS): (1) Treatment Management; (2) Pharmaceutical Requirements; (3) Healthcare Planning and Services and (4) Best Practices Benchmarking.
User-Specific Strategic Knowledge Service Requests: Indeed, different users may need to utilise the healthcare data according to their own requirements that may vary from time to time (subject to the needs at any given time). To incorporate usage and temporal constraints, we allow users to dynamically choose both (a) data items of interest and (b) the data analysis techniques. In this way, we provide user-specific KSHS as opposed to restricting the user to choose from a set of standard services.
Addition of New Strategic Knowledge Services: Our KSHS info-structure caters for the limitless increase in the breadth of knowledge that can be derived from data. This is achieved by adding new data analysis routines to the existing info-structure. The existence of a *data analysis routine library* facilitates new and innovative data analysis routines to be periodically added, as and when needed.
Combination of Multiple Strategic Knowledge Services: Our KSHS info-structure offers the possibility for an 'intelligent' amalgamation of multiple *data analysis routines* in a principled manner to obtain a highly specific data service. In practice, users will be able to dynamically choose pertinent data items from different data analysis routines and systematically synthesise them to yield a unique, yet temporary, data analysis module in line with current needs. This novel concept will elevate our KSHS info-structure to greater acceptability and user satisfaction.
Multiple Data & Knowledge Analysis Techniques: Our KSHS info-structure supports a unique confluence of functionally diverse data analysis methodologies under a single umbrella info-structure. For instance, we will provide data analysis by way of data mining, statistical analysis, rule extraction, time-series forecasting, benchmarking and so on.
Multiple Result Visualisation Methods: Our KSHS info-structure offers a comprehensive and informative view of the data analysed. We offer multiple results visualisation formats—selectable by the user—ranging from graphs to 3D hypercube maps to active reports/documents.

6.0 Concluding Remarks

The HEM project is still under development, in particular the focus of the on-going work is the acquisition of tacit knowledge (Cheah and Abidi, 1999c) and the creation of a variety of medical knowledge bases (Cheah and Abidi, 1999a). Nevertheless the implementation of the KSHS info-structure is ahead of other on-going activities. In fact, we are now able to provide a number of strategic services (on a trial basis), for instance (1) Forecasting 'new diseases' and strategising appropriate preventive measures (cf. Item 6 in Table 2); (2) Forecasting the spread of infectious diseases (cf. Item 8 in Table 2) (Abidi and Goh, 1998a and 1998b); (3) Analysing trends in hospital admission (cf. Item 1 in Table 2)—partially complete; and (4) Analysing treatment pattern (cf. Item 2 in Table 2)—partially complete.

In this paper, we have suggested that the possible synergy of *Knowledge Management* and *KDD* techniques can provide a paradigm for the generation of knowledge-driven strategic services that can go a long way in addressing some of the challenges faced by any modern healthcare enterprise. Our methodology purports the exploitation of experiential knowledge, derived from enterprise-wide data and knowledge bases, for strategic decision-making. The feasibility of our methodology depends on two factors: (a) the availability of a mass of 'knowledge-rich' (healthcare) data with knowledge of healthcare practices and protocols, and (b) the technical capability to extract 'decision-quality' knowledge from data. Firstly, for all practical purposes, modern healthcare systems generate massive amounts of 'knowledge-rich' healthcare data, but unfortunately this asset is not yet fully 'cashed'. Secondly, from a technical point of view, we will like to point out that we have proposed a viable IT info-structure that will help transform healthcare data towards strategic knowledge services. Furthermore, the proposed HEM provides an opportunity to migrate healthcare practice rules, primarily stated in texts, towards the generation of value-added, pro-active strategic services that may directly impact the behaviour and efficacy of the healthcare enterprise as a whole.

In conclusion, we emphasise that the conception of the HEM and the KSHS info-structure has identified opportunities to improve healthcare management through the increased use of cutting-edge information technology. This project has provided us a test-bed to experiment with new strategies and techniques that may evolve from the synergy of knowledge management, knowledge discovery from databases and data mining. Indeed, there have been similar suggestions to combine these two related fields, and for that matter our proposal intends to provide a working illustration of such a strategy. As stated earlier that the project is still in-progress, nevertheless this work undertaken provides us an opportunity to experiment with and analyse the efficacy of KM and data mining techniques in a practical real-life environment.

References

(Abecker *et al.*, 1998) A. Abecker, A. Bernardi, K. Hinkelmann, O. Kühn and M. Sintek. *Toward a Technology for Organizational Memories*. IEEE Intelligent Systems, 13(3): 40-48, 1998.

(Abidi and Goh, 1998a) S.S.R. Abidi and A. Goh. *Applying Knowledge Discovery to Predict Infectious Disease Epidemics*. In H. Lee and H. Motoda (Eds.). Lecture Notes in Artificial Intelligence 1531-PRICAI'98: Topics in Artificial Intelligence, Springer Verlag, Berlin, 1998.

(Abidi and Goh, 1998b) S.S.R. Abidi and A. Goh. *Neural Network Based Forecasting of Bacteria-Antibiotic Interactions for Infectious Disease Control*. In B. Cesnik, A. McCray and J. Scherrer (Eds.). 9th World Congress on Medical Informatics (MedInfo '98): Seoul, IOS Press, 1998.

(Abidi and Yusoff, 1998) S.S.R. Abidi and Z. Yusoff. *Data Driven Healthcare Management: From a Philosophy to an IT Info-Structure*. In International Conference on Multimedia and Information Technology: Kuala Lumpur, August 1998.

(Abidi *et al.*, 1998) S.S.R. Abidi, A. Goh and Z. Yusoff. *Telemedicine and Medical Informatics in the Multimedia Super Corridor: The Malaysian Vision*. In B. Cesnik, A. McCray and J. Scherrer (Eds.). Ninth World Congress on Medical Informatics (MedInfo '98): Seoul, IOS Press, 1998.

(Abidi, 1997) S.S.R. Abidi. *Internet Information Brokering: A Re-Configurable Database Navigation, Data Filter and Export System*. Informatica, 21:185-191, 1997.

(Abidi, 1999) S.S.R. Abidi. *Applying Knowledge Discovery in Healthcare: An Info-Structure for Delivering 'Knowledge-Driven' Strategic Services*. In P. Kokol, B. Zupan, J. Stare, M. Premik and R. Engelbrecht (Eds.). Medical Informatics in Europe (MIE '99), IOS Press, 1999.

(Chandrasekaran and Josephson, 1999) B. Chandrasekaran and J.R. Josephson. *What are Ontologies, and Why Do We Need Them?* IEEE Intelligent Systems and their applications, 14(1):20-26, 1999.

(Cheah and Abidi, 1999a) Y.-N. Cheah and S.S.R. Abidi. *Evaluating the Efficacy of Knowledge Management and Organisational Memory Towards Healthcare Enterprise Modelling*. In Workshop on Knowledge Management and Organizational Memories at Intl. Joint. Conf. on Artificial Intelligence (IJCAI'99): Stockholm, 31 July 1999.

(Cheah and Abidi, 1999b) Y.-N. Cheah and S.S.R. Abidi. *Healthcare Knowledge Management Through Building and Operationalising Healthcare Enterprise Memory*. In P. Kokol, B. Zupan, J. Stare, M. Premik and R. Engelbrecht (Eds.). Medical Informatics in Europe (MIE '99), IOS Press, Amsterdam, 1999.

(Cheah and Abidi, 1999c) Y.-N. Cheah and S.S.R. Abidi. *Tacit Knowledge Creation in a Knowledge Management Context.* In Fourteenth International Symposium on Computer and Information Sciences (ISCIS'99): Kusadasi, Turkey, 18-20 October 1999. (Poster Presentation)

(Davenport *et al.*, 1998) T. Davenport, D. DeLong and M. Beers. *Successful knowledge management projects*. Sloan Management Review, Winter, 43-57, 1998.

(Dorfman, 1997) P. Dorfman. *How Do You Manage Knowledge?* Support Management, 1997.

(Euzenat, 1996) J. Euzenat. *Corporate memory through cooperative creation of knowledge bases and hyper-documents*. Tenth Knowledge Acquisition for Knowledge-Based Systems Workshop (KAW '96): Banff, Alberta, 1996.

(Fayyad *et al.*, 1996) U.M. Fayyad, G.P. Shapiro, P. Smyth, and R. Uthurusamy (Eds.). *Advances in Knowledge Discovery and Data Mining*. AAAI Press, California, 1996.

(Fridsma *et al.*, 1997) D. Fridsma, J. Gennari and M. Musen. *Acquiring, Maintaining, and Customising Organisational Work Process Descriptions*. Proceedings of AAAI Spring Symposium, Artificial Intelligence in Knowledge Management, 1997, pp. 40-46.

(Government of Malaysia, 1997) Government of Malaysia. *CRFP Telemedicine Flagship Application: Lifetime Health Plan*. http://www.mdc.com.my, 1997.

(Gruber, 1993a) T.R. Gruber, *A Translation Approach to Portable Ontologies*. Knowledge Acquisition, 5(2), 1993.

(Gruber, 1993b) T.R. Gruber, *Toward Principles for the Design of Ontologies Used for Knowledge Sharing*. International Workshop on Formal Ontology: Padua, Italy, 1993.

(Kenyon, 1998) J.D. Kenyon. *A Process for Knowledge Acquisition and Management*. Eleventh Workshop on Knowledge Acquisition, Modelling & Management (KAW '98), 1998.

(Kühn and Abecker, 1997) O. Kühn and A. Abecker. *Corporate Memories for Knowledge Management in Industrial Practice: Prospects and Challenges*. J. Universal Computer Science, Special Issue on Information Technology for Knowledge Management, 3(8): 923-954, 1997.

(O'Leary, 1998) D. O'Leary. *Knowledge Management Systems: Converting and Connecting*. IEEE Intelligent Systems, 13(3):30-33, May/June 1998.

(Piatetsky-Shapiro and Frawley, 1991) G. Piatetsky-Shapiro and W. Frawley. *Knowledge Discovery in Databases*. AAAI Press, Menlo Park, California, 1991.

(Simon, 1996) G. Simon. *Knowledge Acquisition and Modelling for Corporate Memory: Lessons Learned from Experience*. Tenth Knowledge Acquisition for Knowledge-Based Systems Workshop (KAW '96): Banff, Alberta, 1996.

(Sveiby, 1997) K. Sveiby. *The New Organizational Wealth: Managing and Measuring Knowledge-Based Assets*. San Francisco: Berrett-Koehler, 1997.

(Turtle, 1995) M.S. Turtle. *Bringing knowledge to the point of care*. Proceeding of Healthcare Information management Systems Society Annual Meeting (HIMSS): Texas, 1995.