In P. Kokol, B. Zupan, J. Stare, M. Premik & R. Engelbrecht (Eds), Medical Informatics in Europe (MIE '99), IOS Press, Amsterdam.

Applying Data Mining in Healthcare: An Info-Structure for Delivering 'Data-Driven' Strategic Services

Syed Sibte Raza ABIDI
Health Informatics Research Group, School of Computer Sciences, Universiti Sains Malaysia, 11800 Penang, Malaysia.

Abstract. Presently, there is a growing demand from the healthcare community to leverage upon and transform the vast quantities of healthcare data into value-added, 'decision-quality' knowledge, vis-à-vis, *strategic knowledge services* oriented towards healthcare management and planning. To meet this end, we present a *Strategic Knowledge Services Info-structure* that leverages on existing healthcare knowledge/data bases to derive decision-quality knowledge—knowledge that is extracted from healthcare data through services akin to *knowledge discovery in databases and data mining*.

1. Introduction

Prime facie, the healthcare enterprise—say hospitals, clinics community centres, etc.—is 'data rich' but 'knowledge poor'. The prevailing scenario is that IT based healthcare systems are consistently churning out volumes of healthcare data; the data may be superficially or partially introspected and exploited by humans for limited and visible intents and purposes; and finally the data is stored in data warehouses either for record keeping or in anticipation of any future use. What is apparent here is that there is a lack of judgement amongst healthcare practitioners with regards to the 'hidden' potential of healthcare data. Thus, there is a strong case for operationalising the seemingly placid healthcare data in to 'health enterprise knowledge', which can then be used for strategic decision making, for instance to abet the planning of healthcare delivery.

The above proposal is not alien to the healthcare industry; one such process that the healthcare industry is well aware of is 'Outcomes Measurement'—a means of searching, or more appropriately 'mining', for previously unknown, actionable information/knowledge from large health databases. By understanding what worked—or did not work—healthcare providers are then able to identify areas for improvement or capitalise on successful methods [1].

In this paper we propose a *Strategic Knowledge Services (SKS) Info-structure* that leverages on existing healthcare data bases to derive decision-quality knowledge—knowledge that is extracted from healthcare data through services akin to *knowledge discovery in databases and data mining* [2] [3]. Here, we will discuss the architectural blueprint of the proposed *SKS Info-structure.* We conclude that the proposed SKS info-structure can add value to routinely collected patient level data and can therefore serve as a management resource to healthcare administrators and policy makers.

2. Introducing Strategic Knowledge Services

Strategic Knowledge Services (SKS) can best be defined as a suite of data-driven strategic services—based on knowledge discovery and data mining techniques—to facilitate the derivation (or extraction) of 'decision-quality' information/knowledge from (a) nation-wide healthcare databases and (b) the health enterprise's knowledge bases, thereby leading to the delivery of value-added, focused and pro-active healthcare services.

Typical SKS may include: data mining for helping health-care providers cut costs and improve care is by showing 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 [4] [5], treatment patterns, hospital admissions, drug patterns and so on; what-if scenario analysis; comparing medical practices (of a segment of practitioners) 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 formulating health policies and planning; data analysis for epidemiological or health surveillance; examining resource allocations and so on [6].

3. Architecture of the SKS Info-Structure

Our proposed SKS info-structure manifests a suite of 'modules' that facilitate the transformation of data to knowledge to wisdom. Conceptually, the SKS info-structure has a multi-tier architecture that effectuates a confluence of knowledge management and data mining techniques. Figure 1 gives an overview of the 4-tier architecture of the proposed SKS info-structure.

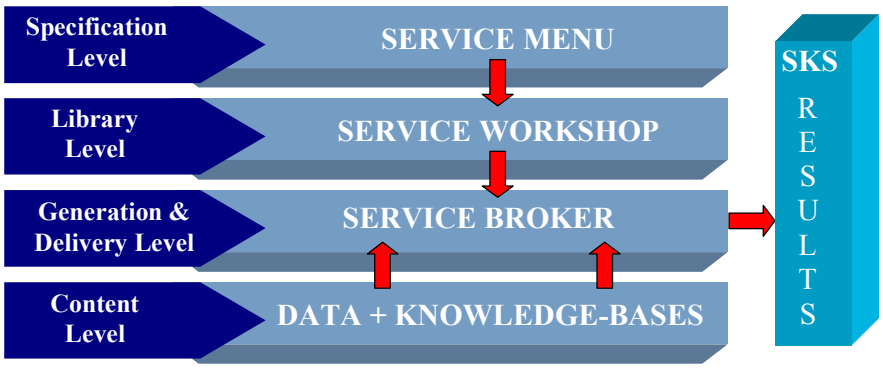

Figure 1: An overview of the multi-tier architecture of the SKS info-structure

In procedural terms, the scheme for the provision SKS is quite simple: (1) specification of the user's desired strategic service(s); (2) procurement of the necessary knowledge and methods (in terms of SKS modules) relevant to the SKS's specification; (3) generating the specified strategic service by the systematic manipulation of the SKS modules in concert with data and knowledge from the data and knowledge repositories; and finally (4) the delivery of the SKS results to users [7].

The SKS 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 SKS. The top level—the *Specification Level*—is responsible for the specification of a particular SKS by users. The second level—the *Library Level*—stores knowledge pertinent to the provision of a variety of SKS; each designated SKS is stored as a specific

SKS 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 SKS 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 SKS 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. Figure 2 gives the architecture of the SKS info-structure.

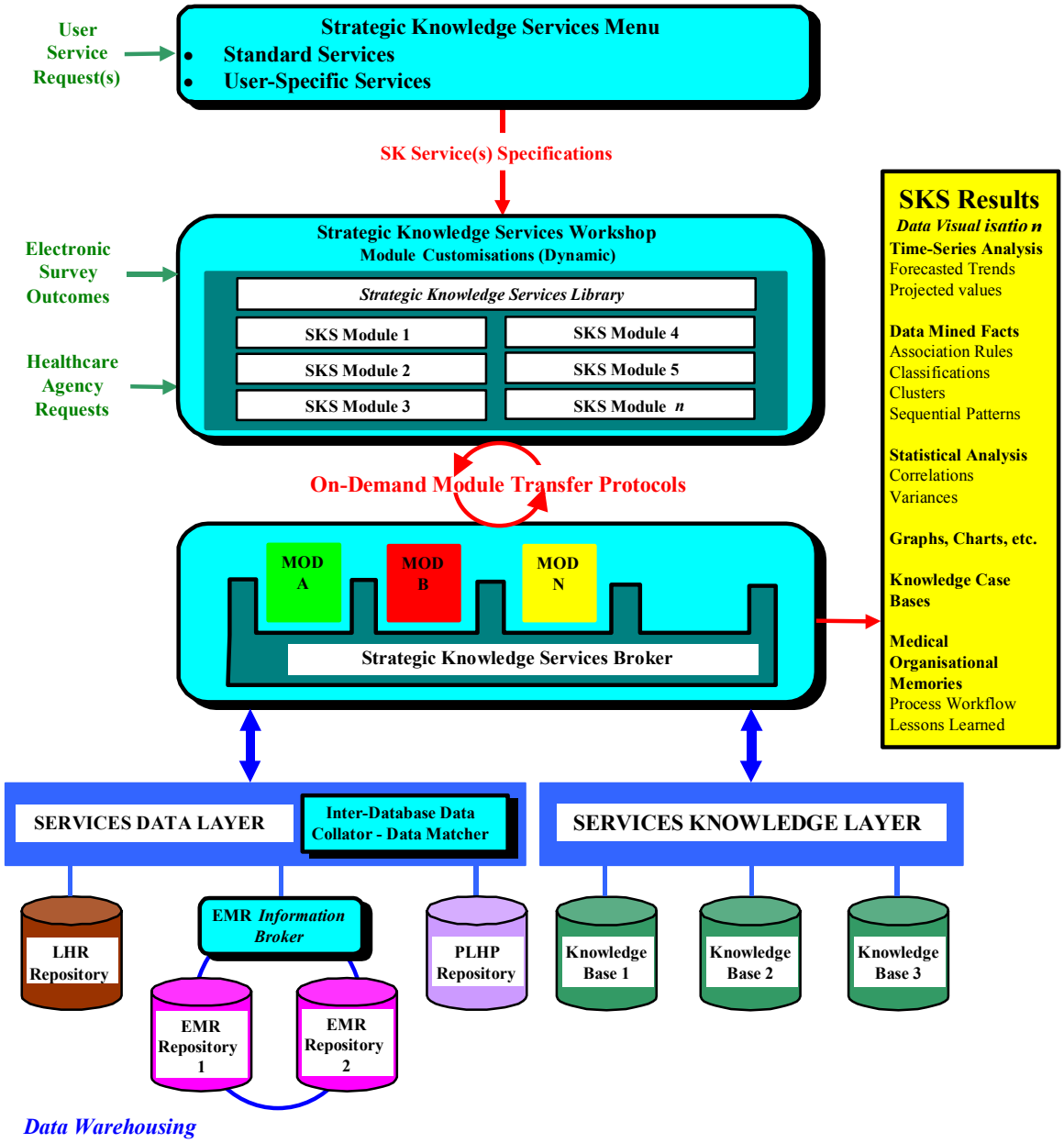

Figure 2: The architecture of the SKS info-structure

4. Features of the SKS Info-Structure

Functionally, the proposed SKS info-structure supports a number of 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 SKS that address the demands of multiple clients—healthcare agencies, insurance and the pharmaceutical industry. The SKS are categorised into four prominent classes (whereas within each class we have numerous sub-SKS): (1) Treatment Management; (2) Pharmaceutical Requirements; (3) Healthcare Planning and Services and (4) Best Practices Benchmarking.

User-Specific Strategic Knowledge Service Requests: To incorporate usage and temporal constraints, we allow users to dynamically choose both (a) data items of interest and (b) the data analysis techniques.

Addition of New Strategic Knowledge Services: Our SKS info-structure caters for the limitless increase in the breadth of knowledge that can be derived from data.

Combination of Multiple Strategic Knowledge Services: Our SKS 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.

Multiple Data & Knowledge Analysis Techniques: Our SKS 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 SKS 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.

5. Concluding Remarks

In this paper, we have suggested that the possible synergy of *Healthcare* and *Data Mining* 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 [8]. Our methodology purports the exploitation of experiential knowledge, derived from enterprise-wide databases, 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.

The implementation of the SKS info-structure is still in its early stages, nevertheless this project provides us an opportunity to experiment with and analyse the efficacy of knowledge management and data mining techniques.

References

[1] A. Weigend & N. Gershenfeld (eds), *Predicting the future and understanding the past*, Addison Wesley, Redwood City, CA, (1993).

[2] U.M. Fayyad, G.P. Shapiro, P. Smyth & R. Uthurusamy (eds.), *Advances in Knowledge Discovery and Data Mining*, AAAI Press, California, (1996).

[3] B. Moxon, Defining Data Mining, DBMS Data Warehouse Supplement, August (1996).

In P. Kokol, B. Zupan, J. Stare, M. Premik & R. Engelbrecht (Eds), Medical Informatics in Europe (MIE '99), IOS Press, Amsterdam.

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

[5] S.S.R. Abidi & A. Goh, Neural Network Based Forecasting of Bacteria-Antibiotic Interactions for Infectious Disease Control". *9th World Congress on Medical Informatics (MedInfo'98)*, Seoul, (1998).

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

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

[8] J. Flower, The Future of Healthcare, Encyclopaedia of the Future, Macmillan & Co., (1995).