

The use of data warehouses in the health-care sector

E. del Hoyo-Barbolla and D. Lees

This paper reviews the use of data warehouses (DWs) in the health sector. It begins by defining what a DW is and considers the reasons for implementing one in a general environment, together with related issues and other concerns. It compares the DW with online transaction process (OLTP) systems.

In the second part, three case studies are presented on how a DW can be applied in the health sector. The three cases are of a different nature: private and public hospitals and a governmental institution. The authors have studied how these institutions have made use of implementing this technical solution in a health environment. Differences between the US and Europe are found when reviewing literature, with the former leading in the commercial health sector. There are, however, important initiatives taking off in the European healthcare sector involving this technology.

In the last section an analysis is carried out on what the benefits and the concerns are when making use of a DW in these environments. The authors' conclusions are that DWs are still not widely used in the health sector and implementing one takes a considerable amount of time, effort and understanding of the impacts of managing one. A DW, is, however, perfectly applicable and can be of great benefit to a healthcare organization as long as confidentiality and security issues are addressed correctly and balanced to get the most out of the DW potential.

Key words

Data warehouse (DW), information management, online transaction process systems (OLTP)

INTRODUCTION

The success of healthcare is based on providing consistent and quality care while finding ways, if possible, of reducing the costs both for patients and organizations.

The healthcare industry is currently focusing on patients as customers. In a competitive climate, if healthcare organizations are to keep their customers, knowing and

managing information about them is essential and organizations realize it is crucial to access viable and timely data.

Integrating data from various sources and turning it into valuable information is a key asset to gain competitive advantage. The concept of DWs fits well with these characteristics: medical, doctor and hospital records including finance information on one hand. On the other, pharmaceutical companies and government records.

The use of DWs in the clinical environment is diverse. With the United States moving towards a 'managed care' environment, DWs are essential to assist the HMO (Health Maintenance Organizations) remain competitive. Other institutions are also implementing them in their transition to a managed care model. In Europe the environment is different. However DWs are still used in hospitals and some central governments.

This paper evaluates some of the issues arising when implementing a DW within a clinical environment. Three different case studies are presented to give insight into the use of DWs.

BACKGROUND: THE DATA WAREHOUSE

Definition

A data warehouse is a subject-oriented, integrated, time-variant and non-volatile collection of data that supports the decision-making process [1] [2].

Some reasons for implementing a data warehouse

The aim of a DW is to integrate organization-wide data into a single repository. Queries can be run, reports and analysis performed to facilitate the decision-making process. A solid platform of integrated and historical data is needed. DW is the centre of the architecture for information systems (IS) in the 1990s.

Online transaction process (OLTP) systems maximized transaction capacity and repetitive, predictable and update intensive business operations. Table 1 summarizes the differences between both systems.

Benefits and problems of the data warehouse

Table 2 synthesizes the benefits and potential drawbacks of implementing a data warehouse. This also applies to the clinical environment.

Eva del Hoyo-Barbolla MSc
DIC MEng Eng (UPM, Spain)
MSc student
34 Holmesdale Road
N6 5TQ, London, UK
Tel: +44 (0)780 181 9146
Email: ehd78105@
GlaxoWellcome.co.uk

David Lees BSc (Hons) PhD
Department of Information
Systems and Computing
Brunel University
Uxbridge
Middlesex UB8 3PH, UK



Fig. 1 Representation of a data warehouse in a company environment [3]

Table 1 Characteristics of OLTP versus DW [1].	
OLTP	DW
Current data	Historic data
Data are dynamic	Data largely static
Repetitive processing	Ad hoc and heuristic processing
Transaction driven	Analysis driven
Application oriented	Subject oriented
Supports day-to-day decisions	Strategic decisions
Serves large number of users	Low number of managerial users

Table 2 Benefits and drawbacks of DW [1].	
Benefits of implementing a DW	Drawbacks when implementing a DW
Potential high returns on investment	Underestimation of resources for data loading
Competitive advantage	Required data non-captured
Increased productivity of decision makers	Data homogenization
Integration of isolated systems	Data ownership
Immediate information delivery	High maintenance needed
Future vision of historical trends	Complexity of integration

DWs require management and marketing. They are good investments only if end users, in our case healthcare professionals, can get at vital information faster and cheaper, achieving better quality transmission than using current technology.

Data may be retrieved from the following sources:

- primary data collected within a department or other sources so research is conducted.
- secondary data acquired to benchmark and monitor performance in the marketplace.
- foreign data from outside the hospital in order to have complete view of the patient and to classify populations.

IMPLEMENTING A DW IN A HEALTHCARE ORGANIZATION: THREE CASE STUDIES

Detroit Medical Centre (DMC)

DMC, the largest healthcare provider in south-east Michigan (eight hospitals and more than 100 outpatient facilities) is, as are many clinical institutions, experiencing dramatic change. In order to survive management realized that access to viable and timely data was essential. A DW project was launched together with a CIS (Clinical Information System) using the same platform. Different systems had co-existed previously, producing conflicting reports in many cases. Clinicians were not able to look at patients across different units. IS were not integrated and difficult and costly to develop and maintain. This certainly represented a challenge from the point of view of information management.

The DW resulted in a 'healthy outcome' [4]. Since its first installation, the vision and goals have incremented significantly. Availability of data is causing more demand for information, ranging from medical record data to purchasing details. Faster and more accurately decisions are made.

The user base increases constantly and it will become a fundamental decision platform for all DMC decision-making levels.

Kaiser Permanente (KP)

KP, the largest HMO in the US, claims to have saved millions of dollars and improved the quality of care with information from its DW. This success is measured by its ability to provide consistent, quality healthcare and by reducing costs. It was first implemented with the goal of obtaining practical measurable results from the management of vast amounts of data.

Several technologies were studied to produce the best results, both medical and non-medical. Technicians and management decided to implement a DW.

Numerous challenges were faced, these being [5]:

- involving management when technology was new and speculative
- training both technical staff and users
- ongoing product evaluation and selection
- few experienced DW practitioners
- enable remote access
- enable a secure access
- multi-dimensionality

They say that 'although the DW is impressive from a technology standpoint, the real innovation is how data can be taken and turned into valuable information that demonstrates how the organization can reduce costs and improve the quality of care'. Benefits already achieved are:

- using 'data mining' for effective drugs: by using the DW it was found that clinicians were using different drugs for the same treatment; protocols have been implemented.
- preventive medicine: guidelines for a particular disease have been developed
- DW has helped in operational expenses and the implementation of best practices.

The Pentagon

The US Department of Defense, and its Health Affairs section (200 hospitals, 400 clinics and other outpatient facilities around the world) decided to move from fee-for-service to a managed care solution. Its Information Systems (IS) team decided to implement a Corporate Executive Information System (CEIS) and a DW for analysis and reporting purposes [6].

Challenges faced included collecting data from the local systems, performing transformations and loading the DW, integrating all data from multiple sources and performing complex operations.

The Pentagon expects to get see benefits such as the capacity to:

- store an enormous amount of data (20 years) and a huge quantity extracted every month
- track the best treatment for various ailments, as well as drugs for illnesses such as malaria or breast cancer.

Many groups are interested in this DW sample knowledge and the Pentagon is currently considering selling part of the data. The biggest concern is the privacy of individual medical records and the potential use done.

Other examples

There are attempts in Europe to develop DW in hospitals and governmental institutions. As an example, the government of The Netherlands strongly believes in the information warehouse for improving healthcare. 'The Amsterdam Experiment' is a DW pilot, the objectives of which are to promote preventive medicine and move to patient-focused care.

New trends are emerging: Web warehousing allows the connection of multiple organi-

zations so as to share information from DW via the Internet. This is still a distant target. Another potential development of importance in hospitals is the integration of all kinds of data: medical images, other electrical medical records, and so on.

Some organizations are moving towards a complete integrated warehouse environment. Standards need to be put in place before this becomes a reality, but efforts are being made in this direction.

Three successful cases have been described. There are many organizations that have found the concept of DW to be extremely useful for many purposes; there are still, however, many issues and concerns to be solved. These are addressed in the next section.

DISCUSSION AND CONCLUSIONS

Implementing and exploiting a DW is not straightforward. Many resources are employed and multiple concerns need to be addressed.

Issues faced in industry apply also in healthcare, but an additional concern affects data management in the clinical environment: security. Being medical data of special sensitivity, it needs to be handled with extreme care.

Understanding the requirements is one of the key challenges. DW will be of no use if it does not suit the staff that actually uses it. The development of an IS requires the participation of the stakeholders, and the extent and effectiveness of the participation is likely to influence the success of the resulting system [7].

Another important challenge is ensuring data quality when integrating data from diverse sources. Some of the problems identified are [1]:

- *Encoding format:* same object is coded differently in different systems.
- *Attribute measurement:* systems have data with identical meaning but use different scale of measurement. This is of special relevance in healthcare.
- *Multiple sources of data:* deciding which is the 'master system', and the criteria used to determine it.
- *Conflicting keys:* DW must have key fields that uniquely identify each record.

Education and training of the user, together with support and documentation when needed, are also concerns in the deployment and maintenance of the DW. Maintenance cannot be left to a separate organization and funding needs to be assured to tackle this. A

possible solution is to establish a cross-functional DW steering committee that would monitor the different stages.

Patients present for treatment expecting that their details will remain confidential. In today's environment, information can easily be assembled and studied. This increased availability raises numerous concerns with respect to privacy. These relate to the danger of unauthorized access and misuse by insiders.

Patients, however, want better health and that requires information sharing. There is a fine line between confidentiality and the need to know.

The Health Insurance Portability and Accountability Act 1996 (HIPAA) requires standards so that information can be shared [8].

Patients can benefit enormously from accessibility of information, when it is 'individually identifiable' (defined as information that identifies a patient or that there is a 'reasonable basis to believe that the information can be used to identify a patient'). Examples of this are:

- research may identify patients within a database that are statistically at risk of contracting a particular disease; preventive medicine can then be used
- access of patients' medical history in emergencies and improvement of their treatment
- searching for potential transplant donors.
- pharmacists can tailor treatments to specific patients
- genetics
- statistical tools applied to large data sets.

The following are examples of how aggregate data (stripped of individual identifiers) can be used:

- to determine seasonal trends to predict supply needs
- for comparison of outcomes of groups of similarly affected patients treated with different drug protocols

- to determine if physicians are performing satisfactorily and saving money.

On the whole, it makes sense to put in place some legal restrictions to avoid abuses in the system. Governments and other institutions need to do more research in order to standardize these practices that concern patients. However, a little loss of privacy may prove to be very healthy for patients and the public in general, and by doing so their lifespan may be increased by some years. The outcome may be worth it.

Both commercial and governmental US institutions are making good use of DW for their purposes. There are still few examples in literature to compare with the European healthcare sector. However, the situation at this side of the Atlantic encourages the adoption of numerous technical solutions to support European organizations. DW are a good example. They will be widely implemented and used in the near future in many organizations including the healthcare industry to the great benefit of their clinical, corporate activity and their patients.

REFERENCES

- [1] Inmon W. Building the data warehouse. New York: John Wiley, 1996.
- [2] Inmon W. What is a data warehouse? Prism, 1995. http://www.cait.wustl.edu/papers/prism/vol1_no1/ (accessed on April 2000).
- [3] Glaxo Wellcome Intranet (accessed on April 2000).
- [4] Feldman J. A dataspace solution for: Detroit Medical Center. *DM Review*, 1999. <http://www.dmreview.com> (accessed on April 2000).
- [5] Hollis J. Deploying an HMO's data warehouse. *Health Management Technology* 1998; **19**(8): 46-48.
- [6] Hemblen M. Pentagon to deploy a huge medical data warehouse. *Computerworld* 1998; **32**(1): 25.
- [7] Cavaye A, Cragg P B. Factors contributing to the success of customer oriented interorganizational systems. *Journal of Strategic Information Systems* 1995; **4**(1), 13-30.
- [8] Mon D, Nunn S. Understanding CPR architecture: a HIM professional's guide. *Journal of AHIMA*, 1999.