DATA WAREHOUSING
a solution for the Caribbean

a Presentation & Proposal

to the **11th Caribbean Conference on**

National Health Financing Initiatives

October 2016   Bonaire, Caribisc Nederland

acknowledgements

This presentation is based on the experience of Ronald Frostestad, Walter Wieners and Raphael Barrett as well as

➤ the work of the *Enterprise Information Systems Steering Committee Data Warehousing Work Group* of the *Healthcare Information and Management Systems Society (HIMSS)* that undertook a review of organisations with successful data warehousing projects in order to develop a model for data warehouse maturity in healthcare

➤ Drew Cardon – Health Catalyst

I would like to acknowledge Ron who passed away on September 14, 2016. Ron had registered to be here.

Data Warehousing – what is it ?

A Data Warehouse is *a centralised repository* that stores data from multiple sources and transforms them into a common, multi-dimensional data model for efficient querying and analysis.

The Data Warehouse enables the collection of existing and new data in support of management's decision-making process.

Data Warehousing – key components

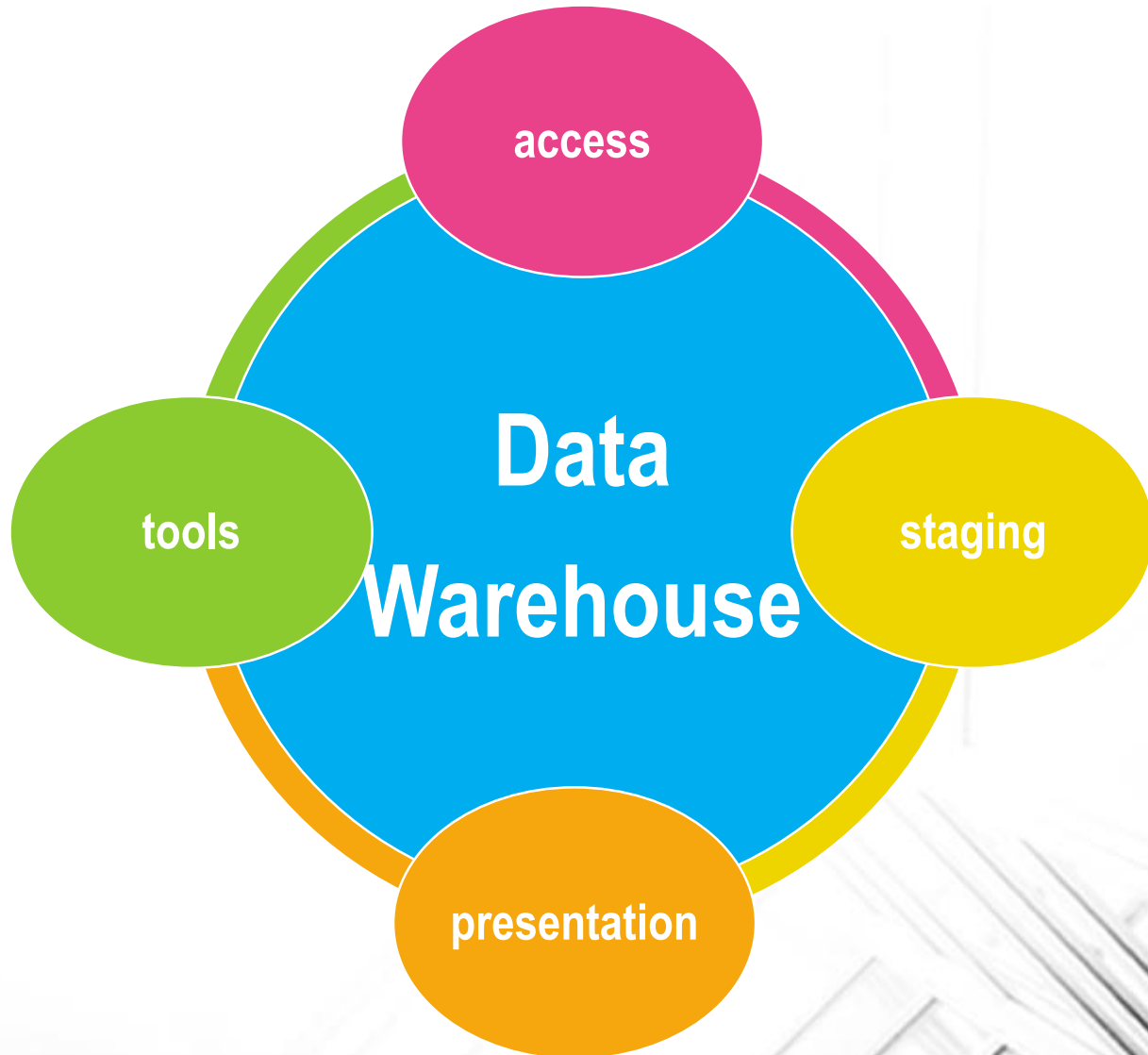Key components of the data warehouse include

On-Line Transaction Processor
- ❖ data access to the operational systems

On-Line Analytical Processor
- ❖ a data staging area
- ❖ a data presentation area
- ❖ data analysis tools

Each component serves a unique function to support the data warehouse

Data access

➤ **On-Line Transaction Processing systems (OLTP) capture and process an organisation's transactional data**

➤ **Typically, each system works independently and shares common data from other operational systems**

➤ **The data acquired through these systems is uploaded to the staging area**

Data staging

The data staging area is both the storage area and the platform to **Extract**, **Transform** and **Load** the raw data into the data warehouse environment.

➢ First, the data is **extracted** from the different sources - operational systems, flat files, manual input, etc.

➢ Then the data is **transformed** - cleansed, formatted and calculated - into a standard format and structure.

➢ Once the data is standardised, it is **loaded** into the presentation area.

Data presentation

The data presentation area is a set of integrated data marts where

➢ A data mart is a subset of the data warehouse with focus on a specific functional area, e.g. finance, clinical, public health

➢ The data in the data mart is relevant to that specific area - detailed and logically organised

Data analysis

Once the data is ready in the presentation area, users can utilise various tools - such as dashboards, scorecards etc. - to access the data for reporting and analysis

Data Warehouse vs Database

A **Database** is typically used for On-Line Transaction Processing (OLTP) constrained to a single application e.g. an electronic health record. They are designed to make transactional systems run efficiently and thus do not readily facilitate analytics.

A **Data Warehouse** uses an On-Line Analytical Processing (OLAP) database to create a layer on top of other databases optimised for and dedicated to analytics taking data from them. The OLAP is structured to perform analytics effectively thus making analytical analysis fast and easy.

Data Warehouse vs Database

	Database	Data Warehouse
Definition	Any collection of data organised for storage, accessibility, and retrieval	A type of database that integrates copies of data from disparate sources for analytical use
Types	On-Line Transaction Processing (OLTP), On-Line Analytical Processing (OLAP), XML, CSV files, flat text and Spreadsheets	An OLAP database that layers on top of other databases to perform analytics
Usage	Typically constrained to a single application: one application equals one database, e.g. EHR	the OLAP provides one source for data that is used for analysis and decision-making

Data Warehouse vs Database

	Database	Data Warehouse
Structure	An OLTP database features very complex tables and joins because the data is *normalised* - no data is duplicated. Making data relational in this way delivers storage and processing efficiencies that allow for fast response times.	An OLAP database facilitates reporting and analysis. The data is *de-normalised* to enhance analytic query response times and ease of use with fewer tables and a structure that results in easier reporting and analyses
Reports	Performing analytical queries is complex usually requiring technical expertise. Reports are static: one-time lists which are helpful for real-time reporting but don't readily allow in-depth analysis.	Various types of analytical reports: Descriptive (*what has happened*) Diagnostic (*why it happened*) Predictive (*what will happen*) Prescriptive (*what to do about it*)

Data Warehouse vs Database

Organisations can harness the power of a Data Warehouse to

- ❖ streamline and scale reporting processes,
- ❖ maintain a single source of truth that everyone can trust,
- ❖ drive meaningful, targeted quality improvement.

By delivering analytics to the frontlines of care as well the boardroom, healthcare organisations can critically evaluate care processes and aggressively pursue the best opportunities for improving outcomes.

Common benefits from implementing a data warehouse were found to include:

- ✓ Reduced costs
- ✓ One consistent data store for reporting, forecasting and analysis – 'the single source of truth'
- ✓ Easier and timely access to data
- ✓ Improved end-user and Information Systems productivity
- ✓ Scalability, Flexibility and Reliability

Data Warehouse advantages

Increasingly, healthcare organisations have used a Data Warehouse to improve clinical, financial and operational decision-making. Here are examples showing the benefits that have been derived.

Texas Children's Hospital in Houston, the largest children's hospital in the USofA, has significantly improved efficiencies with EDW information delivery

➢ *On average, each EDW report costs 70% less to build than an EHR report*
➢ *One EDW visualization replaces 10 EHR reports because the EDW visualizations enable end-users to quickly and easily drill down into the data*

Data Warehouse example: HACs

Hospital-acquired conditions (HACs) cause harm and adversely affect patients' lives whilst increasing hospital length of stay and total costs. **Children's Hospital of Wisconsin,** with two hospitals and a surgery centre, recognised that reducing the rate of HACs required major system and cultural changes with staff education and engagement and a new governance structure powered by high-tech tools and quicker access to new types of data. Implementation of a comprehensive HAC reduction plan resulted in quality of care improvements and cost reductions

➢ *30% reduction in the number of HACs - US$1.6 million savings to-date*

➢ *Individual infection case reductions ranging from 23% to 74%*

Data Warehouse example: H/care delivery

Allina Health is a large integrated healthcare delivery network operating in **Minnesota** and western Wisconsin that includes three large cardiac centres. Due to the high prevalence and mortality rate of CVD, Allina Health recognised that they needed to focus on CVD health in order to truly impact population health and patient outcomes in the communities they serve. Leveraging real-time data from its Data Warehouse, Allina Health effectively identified and addressed clinical practice variation and operational issues affecting CVD care and costs

➢ *>$75 million in performance enhancement savings and revenue increase over a four-year period by focusing on the supply chain, utilisation factors, clinical practice and documentation improvement*

Data Warehouse example: UPMC

The University of Pittsburgh Medical Center (UPMC), with more than 20 hospitals and 500 clinics, recognising that the typical structure of health systems - departments and units within separate hospitals - works against attempts to improve population health and decrease costs determined that it needed to break down the virtual walls between care centres and standardise service lines across the system.

UPMC organised six service lines each spearheaded by clinical, operational, and financial leadership, using the health system's innovative, data-driven activity-based costing methodology to understand the cost of care.

Data Warehouse example: UPMC results

The results of UPMC's service lines and activity-based costing include

- ❖ *$42 million of cost reduction opportunities (2% of service line cost)*
- ❖ *$5 million in supplies savings*
- ❖ *identification of contribution margin variations for specific procedures*
- ❖ *Up to 97% improvement in time to access information*

Women's Health, organised as a service line across the system with strong and collaborative leadership, a Data Warehouse and an activity-based cost management system resulted in the following:

- ❖ *20% reduction in inpatient length of stay for hysterectomies (over 3 years)*
- ❖ *34% reduction in open hysterectomies*
- ❖ *3% reduction in 30-day readmissions for hysterectomies*
- ❖ *25% improvement in contribution margin.*

A Data Warehouse for the Caribbean

All countries in the region have established mechanisms for collecting public health data

- ✓ The data warehouse could be used to improve data collection and analysis for NCCD in addition to communicable diseases such as TB, HIV etc.

- ✓ Better use can be made of this data to help identify trends, hot spots and areas for priority action

- ✓ Established organisations – CARPHA, UWI and PAHO – should be integrated into a regional data warehouse to allow for easier processing, data analysis and research

Caribbean Data Warehouse: INPUTS

INPUT

- ✓ Standardised data collection formats
- ✓ Online submission of raw data to the warehouse
- ✓ Centralised data processing and analysis

For example:

- ➢ Public health field workers could submit data using their mobile phone to a MoH account
- ➢ The MoH would submit this data to the EDW

Caribbean Data Warehouse: OUTPUTS

OUTPUT

✓ **Analyses and Reports**

For example:

➢ **OLTP databases would retain data for each country separately**

➢ **The EDW OLAP would be used for analysis and to generate reports for each country**

➢ **Analysis and reports would also be generated for the region and sub-regions e.g. OECS**

A Planning Toolkit

A Data Warehouse may be part of a larger Toolkit that helps countries improve healthcare delivery and financing.

❖ inexpensive and a-political: may be endorsed by all stakeholders

❖ Useful for other organisations who will be called upon to support the country, e.g. CARPHA, UWI, WHO/PAHO

Better management of data collection and analysis facilitates better estimates and projections of

❖ Population Health   – Disease monitoring

❖ Healthcare Needs   – Public Health efficacy and efficiencies

❖ Healthcare Costs   – Long-term Disease Management & Control