



# IBM数字化医院商业智能解决方案

随着医院信息化在过去十多年的开展，目前大多数大型医院都具备了初步的信息化水平，都拥有了功能完整，强大的HIS,LIS，PACS等业务系统，能够有效支持医院业务的顺利开展，提高了医生的工作效率。运营管理较以前有了很大的进步。

传统的统计报表来自于各个业务系统，信息不全面，同时各个业务系统的数据又存在不一致的情况，这为医院管理者准确掌握医院运行状况带来了困难。因此医院管理者希望实时，准确，快速地掌握医院的运行情况，包括各个业务科室的收入情况，诊疗情况，人员绩效等，以便作出准确的决策。

商业智能解决方案可以帮助客户有效实现以上目标。

## 方案业务需求

决策分析类应用主要是指对有价值的医院数据资源(近期和历史数据)进行集中贮存、汇总、统计和分析的应用。决策分析类应用可用于满足医院管理决策者对医院的经营和运作的监控和决策分析的需求。其主要效益将体现在：通过汇集医院内外的数据，以及合理运用数据分析工具，使医院管理决策层能更方便、更及时地获取关键的管理和业务信息。主要的业务需求包括：

### 财务与管理

- 费用分析
- 成本分析
- 人员绩效考核
- 满意度分析

### 临床

- 临床评价和监督
- 临床效果分析
- 感染控制
- 临床实践分析
- 临床研究

### 质量管理

- 病人安全
- 医疗事故分析
- 用药安全

### 规范遵循

- 循证医疗和规范
- 规范遵循报表

决策分析类应用一般会以数据仓库为基础，然后进一步建设综合统计或者分析型应用。如下图示：



图表 决策分析类应用



数据仓库是一个用于储存企业经营的关键知识及信息的仓库。医院的数据仓库应按不同的业务主题分类，将医院需要从各类业务过程中获取的信息通过标准的集成渠道收集并储存起来。该数据仓库应该可以汇总来自医院资源管理(如财务、人事)、管理支持(如采购、销售等)以及医院运营类应用的数据，实现医院范围内跨科室、跨部门、跨系统的数据集中和整理。

建立数据仓库的目的是根据不同的业务分析需求，将历史数据按照不同的分析维度来进行汇总和呈现，以便为业务决策提供某个维度的信息支持。关于医院的数据仓库，应该囊括医院所有方面的应用系统的数据，并且经过清洗、加工、整合、汇总、转换等过程，形成医院的数据分析、数据展示、动态报表、决策支持以及数据挖掘的统一的数据集中地。

建立数据仓库是一个不断更新的，长时间的循环过程。开始构建时可以提供一些最有价值的信息，实现用户的各取所需，当用户更进一步理解其业务需求和更多的数据需要后，再将更多复杂的数据、基础条件和工具加入数据仓库中。

再者，数据仓库除了存储历史数据以外，也记录了医院业务发展的计划数据和预测数据。所以，借助数据仓库技术，医院的管理人员可以分析经营业绩的完成情况，及产生该结果的原因，从而为业务措施的制定提供帮助。由于数据仓库汇总了财务和非财务、实际和计划、部门和科室的数据，所以将成为医院关键业绩指标的核心内容。

经营分析应用是在数据仓库技术的基础之上，利用数据分析技术对数据进行不同维度的切片处理分析，可得到医院经营情况、收益、投资和汇报等医院绩效评估依据。利用业界标准的具有商业应用价值的统计算法对跨越整个医院的数据进行统计分析，形成对医院现状的统计特征，从而使高层管理者能够从更宏观的角度把握医院的经营状况。例如，现有资源使用情况统计分析、现有服务情况统计分析、供应商/合作伙伴给医院带来的利润统计分析以及医护人员的作业情况比较等等。从功能上讲，经营分析应用涵盖病人帐户情况分析、业务发展分析、收益情况分析、市场竞争分析、服务质量分析、营销管理分析、大客户分析、新业务与现有业务分析、合作服务方分析等方面的主题。

经营分析应用的主要效益体现在：通过汇集医院内外的数据，以及合理运用智能分析工具，使医院决策层、管理层能更方便、更及时地获取关键的管理和业务信息；利用信息挖掘商业价值，辅助决策层对企业的全面监控；通过对内外部数据的分析和利

用，为战略制定、业务拓展、自主定价和运营体系的改善提供决策支持。

## 主要问题和挑战

### 多个异构信息系统带来的信息整合的难度

医院中，通常存在着多个业务系统，典型的三甲医院可能存在几十个业务系统，这些业务系统采用不同的技术开发，运行在不同的软硬件平台上，数据模型差异大，通常缺乏统一的元数据标准，客观上形成了多个信息孤岛。这些都给数据整合带来了极大的困难。

### 数据的数量及其复杂性

医院中的业务系统经过长时间运行，产生了大量的数据，典型的三甲医院的HIS系统的数据量可以达到上百GB。考虑到多个系统的数据整合，数据量通常十分庞大，这就系统的查询性能带来了极大的挑战，需要有专业的技术来处理这些海量数据，例如数据库分区技术，表分区技术，数据压缩技术等，有效提高查询性能和系统管理能力。

这些数据通常存在几个特点：

- 数据之间的关系比较复杂

数据之间存在复杂的关系，需要彻底掌握这种关系，才能有效抽取数据，而不至于发生明显的错误。

- 数据的标准不一致

各个业务系统之间没有统一的元数据规范，在进行数据整合时，就必须要进行数据的格式转换，标准转换等。

- 数据质量存在较大的问题

信息系统的业务数据通常有比较严重的质量问题，例如不符合业务规范要求，明显的错误等，这些数据必须经过数据质量的检查和清洗才可以进入数据仓库系统，否则分析的结果没有任何意义。

### 需要成熟完善的分析方法

用户在对系统数据进行分析的时候，需要多种功能的分析方法以满足不同的需要，例如一般性的报表，灵活查询，图表，仪表盘(dashboard)，多维分析(OLAP)分析，数据挖掘等，此外，还可能需要有非结构化文本分析，嵌入式分析等。



## IBM技术方案

IBM在商业智能领域有着多年的经验，成熟的产品和解决方案，在全球有大量的客户案例。IBM的商业智能方案不是某个单一的产品，也不是几个产品的拼凑，而是为客户提供真正的端到端(end to end)完整的一站式的解决方案，可以满足各个层级的客户需求。

下图是IBM医疗商业智能总体解决方案：



图1 商业智能总体解决方案

IBM有优秀的技术平台，可以帮助客户管理海量的业务数据，提供良好的数据管理能力和多种分析手段。IBM有针对医疗行业的成熟的数据模型，在国外有着大量的成功案例。基于先进的行业模型，我们可以在运营分析，业务分析，绩效考核，医疗质量控制等各方面提供全面完整的主题分析能力，基于IBM在数据仓库项目上的实施方法论，可以帮助客户成功实施商业智能项目，构建完整的解决方案，达成总体业务目标。

IBM解决方案的系统架构如下图所示：

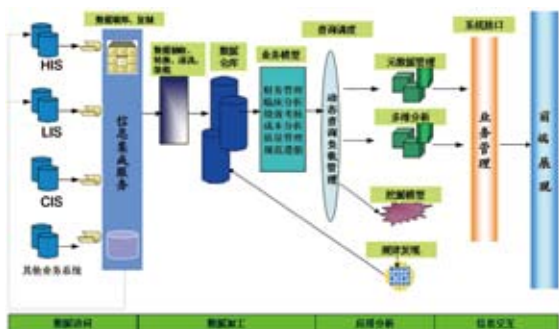


图2 IBM解决方案系统架构

解决方案逻辑上包括4层：

### 数据访问层

该层负责完成信息集成的工作，从各个原有的业务系统中获取业务数据，解决异构数据整合的问题。这是商业智能的第一步。

### 数据加工层

该层负责数据质量的控制，数据清洗，转换，加载等，把数据按照业务要求整理为统一元数据的，符合质量标准的业务数据，并加载到数据仓库平台中。

### 应用分析层

在应用分析层，通过元数据管理，多维分析，数据挖掘等技术，实现对数据的多种分析，满足客户的多层次要求。

### 信息交换层

该层也是所谓的用户访问层，用户通过前端展现工具访问数据仓库平台，以多种灵活的格式展现给客户进行分析研究，例如图表，报表，嵌入式展现等。

## IBM方案中包括的产品

数据集成产品：Information server套件(Federation, Replication)

数据加工产品：DataStage, QualityStage

数据存储产品：DB2 warehouse edition, Tivoli Storage manager, Optim

应用分析产品：Infosphere warehouse

信息交互产品：Alphablox, Cognos

如果想了解IBM单个产品的详细资料，请联系我们的销售代表索取产品手册，或者登陆IBM官方网站下载相关产品介绍资料。

## IBM方案技术优势总结:

### 获取准确完整的信息

实时访问准确且受信任的信息能够帮助企业获取竞争优势。目前，企业IT团队往往面临对业务用户的需求做出迅速响应的挑战，而传统的僵化的数据仓库和孤立的商业智能解决方案成为他们最大的障碍。很多企业都拥有多个数据集市和数据仓库，这可能导致完全不同的、互相冲突的数据。此外，大多数分析解决方案都是与数据仓库分离的，这导致了维护问题，并且需要用户在使用这些解决方案之前将所有数据从数据仓库中复制出来。IT团队需要一种使他们能够支持“单一信息”的业务需求的解决方案，同时必须控制成本、适应不断增长的数据量以及满足越来越多的用户对实时洞察力的需求。

IBM商业智能解决方案是一个完整的、多用途的环境，它允许用户访问、分析和操作任何类型的信息，不管是结构化数据还是非结构化数据、运营数据还是交易数据、实时数据还是历史数据。IBM数据仓库平台提供了用于帮助数据仓库架构师和管理员有效地设计和维护企业数据仓库所需的工具和基础设施。此外，它还是一个开放且基于标准的解决方案，包括数据库、管理和交付工具以及嵌入式分析，以帮助用户创建功能强大、灵活和经济高效的数据仓库解决方案。该解决方案允许用户更深入地分析所有业务数据，更深入地挖掘隐含的关系、模式和趋势，并且通过可靠的、实时的深入理解做出迅速响应。

### 完整的、完全集成的企业数据仓库解决方案

InfoSphere Warehouse是适用于动态数据仓库解决方案的理想平台，因为它提供了远远超越传统数据仓库的强大功能。它的一些关键组件能够帮助您交付动态的业务洞察力，这些组件是完全集成到数据库仓库中的。解决方案包括一组完备的功能和经过验证的IBM技术。

### 功能强大的InfoSphere数据库基础

IBM DB2 9平台是InfoSphere Warehouse解决方案的基础。利用其具有强大可伸缩性的Share nothing分布式架构，DB2 9提供了高性能的混合工作负载查询处理(既可查询关系数据，也可以查询XML数据)。大量高级特性使DB2 9成为一个功能强大的动态数据仓库引擎，这些特性包括数据分区、新的行压缩、多维集群以及物化查询表(MQT)。

### 建模和设计工具

InfoSphere Warehouse 包括基于IBM Rational® Data Architect软件并且能与其进行互操作的建模和设计工具，从而支持用户对物理数据库模式进行设计、建模和实施逆向工程。它提供了一组完备的功能，包括对某个数据库或整个企业进行物理数据建模、模型验证、遵从性分析、影响分析以及变更管理。

### 嵌入式数据移动和转换

InfoSphere Warehouse包含一个用于数据仓库构建和管理的强大图形工具。该工具提供了一组数据流、控制和转换运算符，它们被编译到SQL中，专门针对DB2进行操作且可以在数据仓库中运行。该工具允许用户更有效地准备和填充数据仓库分析结构，以便进行数据挖掘、多维分析和嵌入式分析。

### 数据分区

InfoSphere Warehouse提供了高级的数据分区功能，为用户提供了将数据分发给多个服务器的多种方式，以实现大规模的并行操作和线性的可伸缩性。Share nothing的架构有助于确保性能不会由于数据仓库容量增长而下降。由于InfoSphere Warehouse能够在物理上建立多维数据集，并按照值的范围对数据进行排序，以及将I/O限制到相关的数据分区上，因此它能够极大地减少处理大量查询所需的工作量。此外，按照多个属性对数据进行分组的功能免除了在这些属性上建立索引的需要，从而减少存储需求并降低更新成本。

### 深度压缩

使用InfoSphere Warehouse，用户可以充分利用DB2 9中独创的存储优化技术，它不仅能够极大减少存储关系数据所需的空间和成本，而且还能够提高查询性能。压缩能够节省可观的空间：行业标准TPC-H数据仓库基准测试表明，可节省45%–69%的磁盘空间。现在，压缩功能更易于使用。InfoSphere Warehouse支持用户在将数据装入数据仓库时对数据进行自动压缩，这有助于降低维护成本。

### 极限工作负载管理

IBM InfoSphere Warehouse提供了一组独具特色的极限工作负载管理功能，支持实时交付对业务状况的深入理解，同时继续支持所有传统的信息访问、分析和报告需求，而不会影响性能。



使用传统的数据服务器解决方案很难将信息实时交付给所有用户和应用程序，因为过多的用户运行查询会给底层基础设施带来访问压力，增加并发响应时间。此外，面向大型批处理的数据加载或报告作业，以及由战略和战术规划工作驱动的密集分析处理可能对那些需要快速响应的请求(例如用户与客户进行交互的请求)造成影响。利用InfoSphere Warehouse中提供的高级工作负载管理功能，用户可以为来自不同用户和应用程序的查询划分优先级，并控制分配给这些工作的底层资源总量，从而帮助确保满足关键部门、客户和用户的服务水平协议。

### 嵌入式分析

作为IBM InfoSphere Warehouse软件的一部分，IBM嵌入式分析功能提供了复杂但易于使用的工具集，这些工具被嵌入到数据仓库中。在InfoSphere Warehouse中，IBM已经替您完成了集成工作，因此实现和维护变得更简单，并且提高了用户采用率和您的投资回报。联机分析处理(OLAP)的高级立方体分析、非结构化分析、嵌入式分析以及数据挖掘和可视化功能让您能够充分利用以前未被有效利用的业务信息资源，这些功能紧密协作，构成一个健壮的端到端分析解决方案，为所有用户提供高价值的商业智能。

### 数据挖掘、建模和计分

过去，当要在某个数据集上执行数据挖掘时，必须先从数据仓库中提取数据，然后对其进行单独数据加工和分析处理并将结果发回数据仓库。InfoSphere Warehouse提供了嵌入式数据挖掘、建模和计分功能，允许您直接在数据库中执行分段分析和预测性分析。这意味着您可以使用最新的数据，并实时交付分析结果。InfoSphere Warehouse支持标准的数据挖掘模型算法，例如集群、关联、分类和预测，来自第三方的其他挖掘模型可以以行业标准的“预测模型标记语言”(Predictive Model Markup Language)格式导入到数据仓库中。

### 非结构化分析

大多数商业智能解决方案都无法全面访问在整个企业范围内捕获到的所有信息，例如呼叫中心记录、客户反馈、技术注释和其他自由格式的文本字段，以及文档和Web页面。InfoSphere Warehouse现在提供了分析这些非结构化数据并从先前未充分利用的信息中提取知识的能力。非结构化数据现在可以像结构化数据一样进行访问，从而提供了对客户和产品问题的更深层次的理解。

### OLAP的立方体服务

InfoSphere Warehouse包括OLAP的高级立方体服务，支持多维数据分析，从而允许您将多个业务变量链接到一起，以执行比以前更为深入的分析。特别地，InfoSphere Warehouse立方体服务(cubing service)允许您采用直观的多维和分层导航的方式来分析业务数据，而无需从数据仓库中提取数据。

### 嵌入式分析和可视化

InfoSphere Warehouse包括用于数据挖掘结果的可视分析功能，并且提供了多维分析的显示和导航功能。这些功能是作为丰富的呈现组件提供的，而不通过单独的应用程序提供，因此可以进行轻松的定制并将其直接嵌入到基于Web的应用程序中。通过这种方式，您可以通过用户正在使用的应用程序和门户向他们交付深入的分析结果。

### 利用强大的分析工具来提高灵活性和处理效率

InfoSphere Warehouse Data Studio和Admin Console通过交付公共的集成设计、管理和分析工具简化了数据仓库的分析开发和部署。这些工具为复杂项目的大量设计工作提供便利支持，包括从重新设计数据仓库和填充数据结构，到执行分析以及通过公共界面管理数据挖掘和多维立方体分析等。

InfoSphere Warehouse工具基于开放且可扩展的Eclipse技术构建，并且为构建、测试和调试动态数据仓库元素提供了设计时环境。熟悉的图形开发环境允许用户对数据仓库运算符和实体执行直观的拖放或剪切、粘贴操作。

轻松集成了公共代码管理库(例如IBM Rational ClearCase®软和Concurrent Versions System工具)，支持协作团队开发。基于Web的管理控制台功能为数据库管理员和架构师提供了用于在产品运行时环境中部署和管理InfoSphere Warehouse操作所需的工具。

为帮助用户更快地理解和利用新的环境，InfoSphere Warehouse包括了一个样本数据库环境和手把手的教程，适于辅助加速开发用来提供动态业务洞察力的解决方案。

### 借助DB2 9数据库特性提高数据仓库性能

除了上述高级压缩功能以外，DB2 9 MQT和多维集群也有助于提高性能。MQT是一种经过预先汇总和预先聚合的表，它将查询结果保存为数据。DB2软件优化器能够透明地将来自基本表的查询重新定向到匹配的MQT，从而极大地提高复杂聚合查询的性能。DB2中的普通表和MQT可以跨多维集群进行组织，以帮助优化磁盘I/O和加载/卸载操作，从而提高分析查询性能。

