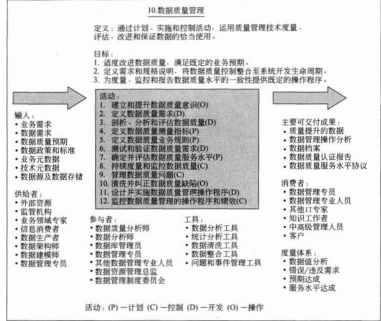


数据质量管理

基础



关联图

数据质量管理

计划-实施-检查-行动

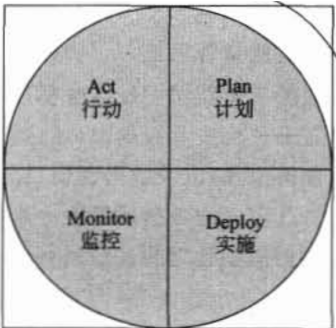


图 12.2 数据质量管理循环

- 制定数据质量管理评估计划和识别数据质量管理关键目标。
- 实施数据质量管理评估计划。
- 监控和度量数据质量管理水平，以提供数据质量从可更好满足业务需求。

步骤描述

定义数据质量需求

数据质量维度

- 准确性
- 完整性
- 一致性
- 时效性
- 隐私性
- 合理性
- 参照完整性
- 及时性
- 唯一性
- 有效性

实现数据质量评估

步骤

- 确定评估数据集
- 记录业务使用方式
- 利用数据剖析工具和技术对数据分析
- 列出全部的潜在异常
- 对每个异常和专家进行评估，并分析潜在的影响
- 对重要的数据排序，定义评价指标

定义数据质量评价指标

- 可度量性
- 业务相关性
- 可接受程度
- 数据管理制度
- 可控性
- 可追踪性

定义数据质量业务规则

基于数据对业务的影响程度

- 对不满足业务的数据记录
- 生成通知事件，提示潜在的数据质量问题
- 建立自动的或是事件驱动的缺陷数据纠正机制，满足业务需求

模板定义

- 值域成员
- 定义一致性
- 值域的一致性
- 格式一致性
- 映射一致性
- 取值和记录的完整性
- 一致性规则
- 精确度验证
- 唯一性验证
- 及时性验证

数据质量工具

- 数据剖析
- 解析和标准化
- 数据转换
- 身份识别和匹配