Executive Summary/Abstract
Data Diagnosis: a practical “RAD” approach to data quality from several years working with tricky telecommunications data. The lessons learned are applicable across any data in any sector. This presentation shares how these tactics can address gaps in data knowledge, ensure early wins, at the same time underpinning a “TQM” approach to data quality across an organization.

Objectives of this presentation
- Set the scene – the “Telco” origins
- Introduce data knowledge, it’s significance and how TDQM helps
- Position data discovery and profiling
- Introduce the data diagnosis framework
- Discuss “Low Hanging Fruit” payback opportunity(s)
- Align data diagnosis to competitive advantage
- Show data diagnosis underpinning Information Quality system(s)

Outline and Structure
- Brief overview of typical global telecommunications delivery processes and data resource requirements
- The typical Telecommunication’s provisioning information chain
- The intricacies of telecommunication’s data information chain
- “Real World” – the data quality risks experienced

Tier 1 Telco’s service operation

Outline centric data is tricky!!
Network data domain .... the reality

- Missing or imprecise data integrity rules are common due to difficult data structures, lack of domain expertise and poor documentation
- Data error detection and correction is expensive and tricky, weighed down by inadequate metadata
- Everyone recognizes meta data shortcomings, just don’t know what to do about it, it’s not uncommon for meta data to be missing altogether
- Only a fraction of data created automatically is a) used and b) important
- Many major data quality problems are down to poor training and lack of validation, they are relatively easy to fix.

How Data Quality “Best practice” can be brought to bear

- Data Diagnosis within TDQM based Information System
- What does Data Quality mean
- How to make Data Quality Assessment work

The impact of Data Diagnosis on an IQ System [3]

The challenge .... “Fit for purpose” data health check [1]

Information Quality - Defined

Data are of high quality if they are fit for their intended uses in operations, decision-making, and planning (after Juran).

- Accuracy
- Integrity
- Correctness
- Completeness
- Consistency
- ... etc.

Importance and worth

- Free from defects
  - accuracy
  - integrity
  - correctness
  - completeness
  - consistency
  - ... etc.

- Possess desired features
  - comprehensive
  - relevant
  - current
  - appropriate
  - ... etc

Exploiting the TDQM CYCLE to enhance Data Quality Assessment [2]

The data quality assessment process followed

- A Data Discovery phase to build Data resource knowledge
- Data Diagnosis as a systematic data content test and scoring mechanism
- The resulting Data Quality scorecard
Data Discovery kick starts Data Knowledge

Data Diagnosis implements incremental IQ assessment

Data Quality Scorecard ...... Measurement replaces debates

Data Discovery kick starts Data Knowledge

Data Diagnosis ...... exploits “low hanging fruit” first

Data Diagnosis ...... achieves results

Data Diagnosis implements incremental IQ assessment

The typical detail result and business benefit

- The type of data integrity error, its impact and financial impact possible
- The business case for formal data assessment and TDQM

Data Discovery kick starts Data Knowledge

Data Diagnosis implements incremental IQ assessment

Data Quality Scorecard ...... Measurement replaces debates

Data Diagnosis ...... exploits “low hanging fruit” first

Data Diagnosis ...... achieves results

Data Discovery kick starts Data Knowledge

Data Diagnosis implements incremental IQ assessment

The typical detail result and business benefit

- The type of data integrity error, its impact and financial impact possible
- The business case for formal data assessment and TDQM

Data Discovery kick starts Data Knowledge

Data Diagnosis implements incremental IQ assessment

Data Quality Scorecard ...... Measurement replaces debates

Data Discovery kick starts Data Knowledge

Data Diagnosis implements incremental IQ assessment

The typical detail result and business benefit

- The type of data integrity error, its impact and financial impact possible
- The business case for formal data assessment and TDQM

Data Discovery kick starts Data Knowledge

Data Diagnosis implements incremental IQ assessment

Data Quality Scorecard ...... Measurement replaces debates

Data Discovery kick starts Data Knowledge

Data Diagnosis implements incremental IQ assessment

The typical detail result and business benefit

- The type of data integrity error, its impact and financial impact possible
- The business case for formal data assessment and TDQM

8th International Conference on Information Quality, 2003

Proceedings of the Eighth International Conference on Information Quality (ICIQ-03)

154
References