Abstract

In recent years, corporate scandals, regulatory changes, and the collapse of major financial institutions have brought much warranted attention to the quality of enterprise information. If we understand the underlying sources of quality issues, then we can develop a plan of action to address the problem that is both proactive and strategic. The relationship between data quality and data consistency is investigated in this research paper. Poor data quality is costly as it adds expenses and lowers user satisfaction so Data Profiling could be enforced. It was observed that quality of data in a data warehouse is affected by factors like: data not fully captured, lack of planning, data aging.
References

- Scott W. Ambler (2001) Challenges with legacy data: knowing your data enemy is the first step in overcoming it, practice Leader, Agile Development, Rational Methods Group, IBM, 01 Jul 2001
- Clements, R. B., 1990, Creating and Assuring Quality, ASQC Quality press
- Wuisdom, J. "Research Problems in Data Warehousing"; 4th International CIKM Conf. 1995
- J. Dean and S. Ghemawat. MapReduce: Simplified Data Processing on Large Clusters. In OSDI, 2004

Index Terms

Computer Science
Data Mining

Keywords
Aging  Data Consistency  Data Profiling  Data Quality  Data Warehouse.