

Design and Implementation of Data Quality Measurement Tool

**A dissertation submitted to the University of Manchester
for the degree of Master of Science in the Faculty of Engineering and
Physical Sciences**

2012

Michalis Nikolaou

School of Computer Science

Abstract

“Garbage in garbage out” a well-known phrase used to emphasize the fact that when there is poor Data Quality in an information system, such as a database, then do not expect to get useful information out of such a system. Questions that arise though are: How many enterprises understand what Data Quality means? How many enterprises are aware that they preserve poor Data Quality? How many enterprises are aware about the presence of Data Quality measurement and improvement tools?

In the world of Business, very few enterprises understand the concept of Data Quality and much less try to improve the quality of their data centers. As Olson said, “It takes data to change their minds” [Olson, 2003] referring to enterprises and businessmen.

On the other hand, Data Quality is a widely acknowledged concept in the world of Information Technology. Nowadays, experts do not deal only with Data Quality as a theoretical concept but they have moved beyond that, to design assurance tools and programs in order to solve actual Data Quality business problems.

The proposed thesis suggests the development of a practical Data Quality measurement tool. Its measurements are based on four Data Quality dimensions: Accuracy, Consistency, Timeliness and Completeness – most of the problems encountered regarding Data Quality are under these dimensions. It is a web application combined with an easy to use User Interface (UI) that does not require any advance technical knowledge. It mainly addresses IT oriented people but small enterprises with no in-house IT experts can use it to assess their Data Quality.