Chapter V

Applying Name Knowledge to Information Quality Assessments

Kimberly D. Hess, CASA 20th Judicial District, USA
John R. Talburt, University of Arkansas at Little Rock, USA

ABSTRACT

An introduction to name knowledge and its application to information quality assessments through an expert system is discussed in this chapter. The quality of name information has become an increasingly important issue as companies strive to implement customer relationship management (CRM) strategies in which the customer name plays an important role in the entity resolution process for data integration applications — ultimately impacting customer recognition systems. As many applications have been developed and refined for assessing and improving the quality of mailing address information, the potential exists to affect a similar success for customer name information. This chapter discusses both theoretical and practical considerations in the approach, design, and administration of systems for assessing the quality of name information.

INTRODUCTION

As companies worldwide attempt to implement Customer Relationship Management (CRM) strategies, they begin to understand their success is directly dependent upon the quality of the data (Huang, 1999). For this reason, these same
companies are also adopting information quality management practices, such as Total Data Quality Management (TDQM) (Wang, 1998) and Six Sigma. These practices are guidelines for approaching and implementing measurable quality initiatives. Necessary requirements for improving information quality are to establish desired levels of quality (goals), measure actual quality, analyze failures, and implement improvements. In the case of CRM, name information quality along with address information quality can be seen as the foundation for building enduring customer-centric relationships.

Although companies often maintain many items of information about their customers, the poor quality of the core elements of name and address are often the roadblocks to the successful and cost-effective data integration. Traditionally, processes have focused on the address portion of customer contact information, motivated primarily by postage savings realized from the reduction of undeliverable mail and U.S. Postal Service discounts offered for presorting large mailings. Often processes that extracted address information from raw transaction data would simply aggregate and move any residual information into a name field with little or no name information quality processing. Organizations are now realizing that the name components must also be analyzed and improved to ensure accurate Customer Data Integration (CDI).

The value of an enterprise’s existing expertise on name information can be drawn from a variety of resources and leveraged throughout the enterprise by embedding it into knowledge-based-driven expert systems. These systems usually employ some combination of software capabilities, encompassing standard and comprehensive name knowledge reference tables (e.g., common names, vulgar names, standard abbreviations) and pattern analysis. If for no other reason, consolidating name knowledge across an enterprise into a standard assessment tool can be a huge step toward improving the consistency of name information and lessen the complexity of algorithmic resolution of name information in downstream processing. From a standard assessment, additional business rules can be developed and applied for determining the acceptable level of name information quality allowable in specific processing steps or applications. The knowledge exercised in the evaluation of name data utilizing an expert system can provide significant improvement in the overall results of processing from data integration to complex recognition solutions.

BACKGROUND

When the Customer Data Management Industry was primarily focused on Direct Mail Marketing (DMM), postal delivery point validation, cleansing, and correction were elevated to a fine art, but minimal quality checks were applied to the name component — a mail piece addressed to “RESIDENT, 123 OAK STREET” was just as deliverable as one sent to “JOHN SMITH, 123 OAK STREET.” With the current emphasis on Customer Relationship Management (CRM) and its dependence
Quality Metrics and Bayesian Analysis: The Case of Extreme Programming
www.igi-global.com/chapter/quality-metrics-bayesian-analysis/26167?camid=4v1a