

Chapter 3

Subjective Information Quality in Data Integration: Evaluation and Principles

Ahmed AbuHalimeh

University of Arkansas – Little Rock, USA

M. Eduard Tudoreanu

University of Arkansas – Little Rock, USA

ABSTRACT

This chapter focuses on the science of human perception of information quality and describes a subset of Information Quality (IQ) dimensions, which are termed Subjective Information Quality (SIQ). These dimensions typically require a user's opinion and do not have a clear mathematical technique for finding their value. Note that most dimensions can be measured through multiple techniques, but the SIQ ones are most useful when the user's experience, opinion, or performance is accounted for. This chapter explores SIQ while considering information obtained from multiple sources, which is a common occurrence when employing visualizations to perform business or intelligence analytics. Thus, the issues addressed here are the assessment of subjective perception of quality of data shown through visual means and principles on how to estimate the subjective quality of combined information sources.

INTRODUCTION

Modern analytics tools employed in business, intelligence, and even science, rely increasingly on rapid exploration of multiple information spaces and on a combination of various information dimensions and sources performed by an analyst.

DOI: 10.4018/978-1-4666-4892-0.ch003

Integrating information quality involves providing users with a unified view of this data quality. This process becomes significant in a variety of situations both commercial and scientific, particularly with the widespread use of information visualization to allow end-users prompt access to large amounts of data. Integrating information appears with increasing frequency as the volume and the need to share existing information grows.

Information quality (IQ) is not a simple scalar measure, but can be defined on multiple dimensions, with each dimension yielding different meanings to different information consumers and processes. Each dimension can be measured and assessed differently. Information Quality assessment implies providing a value for each dimension about how much of the dimension or quality feature is achieved in order to enable adequate understanding and management.

This chapter focuses on a subset of IQ dimensions, which we term subjective information quality (SIQ). These dimensions typically require a user's opinion and do not have a clear mathematical technique for finding their value. Note that most dimensions can be measured through multiple

techniques, but the SIQ ones are most useful when the user's experience, opinion, or performance, are accounted for. The objective dimensions are the dimensions that can be largely and commonly assessed via mathematical or functional forms. It is possible for the same dimension to be measured by both subjective and objective means, depending on the context. For example, accuracy can be objectively computed when one checks the contents of a warehouse against an electronic inventory record. However, accuracy can also be subjectively employed, as for example, the estimation of the amount of shade in a specific area. Table 1 shows all IQ dimensions and the relative importance or usefulness of objective metrics versus subjective ones (Pipino, Lee, & Wang, 2002).

Table 1. IQ dimensions and the relative usefulness of objective metrics versus subjective ones

Dimension	Objective Assessment (width of gray area proportional to relevance of objective metrics)	Subjective Assessment (width of white area shows relevance for subjective)
Believability	Apply a formula (integration only)	User's opinion and experience determines whether they trust the data
Ease of Manipulation	Time to perform a computation	User's experience or performance with the data
Interpretability	Whether some computation is successful	User can understand the data correctly
Relevancy	Can produce a valid result	Helps the user in their task
Reputation	Apply a formula (integration only)	The user can judge or assume the accuracy based on the result of the objective assessment
Value-Added	Can increase the value of data	The user can judge or assess the value added to the data
Objectivity	Formulas applied	User opinion
Timeliness	Can reflect how up-to-date the data is with respect to the task	User judges based on previous experience
Understandability	Can provide clear and simple data	User can understand the data easily
Concise Representation	The shortest representation is known	User judges based on previous experience
Appropriate Amount of Data	The needed amount is known	User expertise is required
Security	Against a standard metric	Users experience or performance with the data
Accessibility	Against a standard metric	Based on user's experiences
Consistent Representation	Count different representations	User's opinion
Accuracy	Formula based on known, exact value	Expert estimation when exact value not available
Completeness	Count missing values in structured sources	User's opinion for unstructured text

20 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/subjective-information-quality-in-data-integration/96144

Related Content

Measuring the Effect of EWOM Readability and Sentiment on Sales: Online Cellphone Reviews

Aakash Aakash and Anu Gupta Aggarwal (2020). *International Journal of Business Analytics* (pp. 24-42).

www.irma-international.org/article/measuring-the-effect-of-ewom-readability-and-sentiment-on-sales/264261

Supervised Regression Clustering: A Case Study for Fashion Products

Ali Fallah Tehrani and Diane Ahrens (2016). *International Journal of Business Analytics* (pp. 21-40).

www.irma-international.org/article/supervised-regression-clustering/165009

Speed of Technology Adaptation in Connection to Organizational Change and Ownership Concentration: Study in Croatia

Lovorka Galetic, Najla Podrug and Domagoj Hruska (2010). *Pervasive Computing for Business: Trends and Applications* (pp. 130-149).

www.irma-international.org/chapter/speed-technology-adaptation-connection-organizational/41101

Profit Maximization Modeling for Supply Chain Planning

Tan Miller and Renato de Matta (2014). *Encyclopedia of Business Analytics and Optimization* (pp. 1910-1921).

www.irma-international.org/chapter/profit-maximization-modeling-for-supply-chain-planning/107379

Business Intelligence Conceptual Model

Fletcher H. Glancy and Surya B. Yadav (2011). *International Journal of Business Intelligence Research* (pp. 48-66).

www.irma-international.org/article/business-intelligence-conceptual-model/53868