Journal of Theoretical and Applied Electronic Commerce Research
ISSN 0718–1876 Electronic Version
VOL 9 / ISSUE 2 / MAY 2014 / 64-79

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762014000200006

A Metrics-Driven Approach for Quality Assessment of Linked Open Data

Behshid Behkamal[1], Mohsen Kahani[2], Ebrahim Bagheri[3] and Zoran Jeremic[4]

[1] Ferdowsi University of Mashhad, Computer Engineering Department, Mashhad, Iran, Behkamal@stu.um.ac.ir
[2] Ferdowsi University of Mashhad, Computer Engineering Department, Mashhad, Iran, Kahani@um.ac.ir
[3] Ryerson University, Department of Electrical and Computer Engineering, Toronto, Canada, bagheri@ryerson.ca
[4] Ryerson University, Department of Electrical and Computer Engineering, Toronto, Canada, zoran.jeremic@ryerson.ca

Received 2 August 2013; received in revised form 1 November 2013; accepted 11 December 2013

Abstract

The main objective of the Web of Data paradigm is to crystallize knowledge through the interlinking of already existing but dispersed data. The usefulness of the developed knowledge depends strongly on the quality of the published data. Researchers have observed many deficiencies with regard to the quality of Linked Open Data. The first step towards improving the quality of data released as a part of the Linked Open Data Cloud is to develop tools for measuring the quality of such data. To this end, the main objective of this paper is to propose and validate a set of metrics for evaluating the inherent quality characteristics of a dataset before it is released to the Linked Open Data Cloud. These inherent characteristics are semantic accuracy, syntactic accuracy, uniqueness, completeness and consistency. We follow the Goal-Question-Metric approach to propose various metrics for each of these five quality characteristics. We provide both theoretical validation and empirical observation of the behavior of the proposed metrics in this paper. The proposed set of metrics establishes a starting point for a systematic inherent quality analysis of open datasets.

Keywords: Metrics, Linked open data, Correctness, Consistency, Quality assessment

Journal of Theoretical and Applied Electronic Commerce Research
ISSN 0718–1876 Electronic Version
VOL 9 / ISSUE 2 / MAY 2014 / 64-79

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762014000200006

1 Introduction

Linked Open Data (LOD) or the Web of Data refers to a set of best practices for publishing and connecting structured data, which allows data providers to publish their publicly available data and meaningfully link them with other information sources over the Web. The main goal of the LOD initiative is to create knowledge by interlinking dispersed but related data instead of linking related documents in the traditional Web. The three key technologies which support LOD include URI (Uniform Resource Identifier), a means to identify entities or concepts, HTTP (Hyper Text Transfer Protocol), a simple mechanism for retrieving resources, and RDF (Resource Description Framework), a data model for describing and linking data. The current LOD cloud consists of over 50 billion facts represented as RDF triples covering a diverse set of domains such as health, government, and music. This massive amount of data on the LOD opens up significant challenges with regard to data quality. For example, datasets that contain information extracted from semi-structured, or even unstructured sources, such as those presented in DBpedia [34], often contain inconsistencies as well as misrepresented and incomplete information. Given the fact that it is the quality of data that determines its usefulness and real applicability, we believe that the success of the LOD initiative depends strongly on the quality of the datasets that are released to the LOD cloud.

In recent years, researchers have developed several methodologies, metrics and tools to evaluate data quality in general. For example, [28] describes subjective and objective assessments of data quality and presents three functional forms for developing objective data quality metrics including simple ratio, min/max operation and weighted average. In [23], the authors have proposed a methodology for the assessment of organizational Information Quality (IQ), which consists of a systematic questionnaire to measure IQ and is accompanied by analysis techniques for interpreting the proposed IQ measures. In the area of the methodologies for data quality assessment, [3] provides a comparative description of existing methodologies and provides a comprehensive comparison of these methodologies along several dimensions, including the methodological phases and steps, the strategies and techniques, the data quality dimensions, the types of data, and, finally, the types of information systems addressed by each methodology. Also, the database community has developed a number of approaches such as user experience, expert judgment, sampling, parsing and cleansing techniques [4], [26] for measuring and enhancing data quality. While such approaches serve as the guiding background knowledge for the measurement of data quality on the LOD, they cannot be directly applied. This is in essence due to the fact that data quality on the LOD is related to novel aspects such as data representation quality or consistency with regard to the information presented in the other published datasets. In addition, mechanisms of knowledge inference on the LOD usually follow an open world assumption, while the mentioned methods usually adopt closed world semantics.

While data quality is an important requirement for the successful organic growth of the LOD, only a very limited number of research initiatives exist, which focus on data quality for the Semantic Web and specifically for LOD. Based on our practical experience in publishing linked data [6], we have observed that many of the published datasets suffer from quality issues such as syntax errors, redundant instances, and incorrect/incomplete attribute values. One of the better strategies to avoid such issues is to evaluate the quality of a dataset before it is published on the LOD cloud. This will help publishers to filter out low-quality data based on the quality assessment results, which in turn enables data consumers to make better and more informed decisions when using the shared datasets. Based on the model proposed by the W3C's Linked Data Cookbook [5], there are seven essential steps for publishing any data as LOD. These steps are:

1. Model the data
2. Name things with URIs
3. Re-use vocabularies whenever possible
4. Publish human and machine readable descriptions
5. Convert data to RDF
6. Specify an appropriate license
7. Host the linked dataset publicly and announce it

Since our work focuses on the assessment of the quality aspects of a dataset itself and not the context of other datasets; the work that we propose should be applied on data before their release as a part of the LOD cloud. According to this model, the best time to perform the assessment is after converting source data into RDF which is presented in the fifth step of this model. As a result, we first need to identify inherent quality characteristics that focus only on the quality of data sources. For this purpose, we will explore the characteristics of data published on the LOD cloud to extract the inherent quality characteristics and propose a set of metrics, which are quantitatively measurable for a given dataset. This way, we are able to assess inherent quality characteristics of datasets before publishing the

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762014000200006

data by observing the measured values of the relevant metrics. Based on our approach, the main research question that we will be trying to answer is the following: *Can a set of measurement-theoretic metrics be defined to assess the inherent quality of a dataset?* The idea behind this research question is derived from the area of software quality measurement, where metrics are defined as measurable syntactic aspects of software artifacts, such as Lines of Code, Cyclomatic Complexity and Inheritance Depth. The novel contributions of our work can be summarized as follows:

- We clearly identify a set of important inherent quality characteristics for LOD datasets based on existing standard quality models and frameworks (such as ISO-25012).
- We systematically propose and validate a set of metrics for measuring the quality characteristics of datasets before they are published to the LOD cloud.
- We propose a quality model for LOD that considers the inherent data quality indicators of such data.
- We introduce a novel approach for the assessment of the quality of datasets on LOD, which has its roots in measurement theory and software measurement techniques.

The rest of this paper is organized as follows: Section 2 reviews existing data quality approaches in the LOD domain. Section 3 discusses the quality characteristics which depict our notion of inherent quality of data in the context of LOD. Our proposed approach for measuring inherent quality of LOD datasets developed based on the Goal-Question-Metric Approach is introduced in Section 4. Evaluation of the work and some general discussions are provided in Section 5. Finally, the paper is concluded in Section 6.

2 The State of the Art

In this section, we primarily focus on data quality with respect to LOD. However we also briefly touch upon other quality assessment frameworks that are relevant to our work.

Despite its importance, data quality has only recently being receiving attention from the Semantic Web community. For this reason, there are only a limited number of works, which specifically focus on the quality of data in LOD. The authors in [19] propose a framework to assess the information quality of Web data sources based on provenance information. Furthermore, Bizer and Cyganiak have developed a framework, called WIQA, which filters poor information in Web-based information systems according to user defined quality requirements [8]. Other approaches have used Semantic Web technologies to identify and correct data quality issues. For instance, the approach proposed by [24] has exploited both domain and background knowledge to detect data deficiencies in metadata including spurious annotations, data inconsistencies and inaccurate data. A more recent work has used ontologies to annotate incorrect data, such as redundant instances or incorrect attribute value combinations to train detection algorithms for automated identification of data quality problems in cancer registries and data sources from the energy industry [12].

The related literature in the context of LOD can be classified into two main groups. The first group of related work investigates the quality problems of the published datasets. The authors of [20] have proposed a comprehensive approach that classifies quality problems of the published linked datasets and discusses common errors in RDF publishing, their consequences for applications, along with possible publisher-oriented approaches to improve the quality of machine-readable and open data on the Web. In another work, Furber and Hepp proposed an approach to evaluate the quality of datasets using SPARQL (Simple Protocol and RDF Query Language) queries in order to identify quality problems such as missing literal values or data type properties, illegal literal values, and functional dependency violations. Using this approach, the authors identify quality problems of already available datasets such as Geonames and DBPedia [17]. The second group of related works includes some tools for checking RDF documents, each with its own error-checking functionalities. Some, which are available online, accept an RDF/XML (Resource Description Framework/eXtensible Markup Language) document as input and check whether the document is syntactically valid or not. Other kinds of online validators, such as URIDebugger [15] and Vapour (Site 1), check the dereferencability of a given URI and determine whether the given URI is an information resource or a non-information resource. Other platforms, such as Jena Eyeball (Site 2) and VRP [32] which are often in the form of command line tools, are designed for identifying common errors in RDF documents. Generally, all of these works primarily focus on data quality problems in published datasets, and none of them provide a concrete solution for improving data quality, nor attempt to identify the causes of the quality problems before the data is published. Moreover, none of the above approaches propose a method or a set of metrics for systematically evaluating inherent data quality.

In this paper, we argue the importance of filtering out poor quality data by assessing the quality of a given dataset before publishing it to the LOD cloud. Furthermore, we propose a methodology to identify quality-related problems and suggest how we may systematically evaluate inherent data quality. To this end, our approach for data quality assessment involves the measurement of quality dimensions or criteria that are focused specifically on inherent

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762014000200006

aspects of data quality. We propose a metrics-driven approach for assessing the quality of data to be published as a part of the LOD cloud. We accomplish this by identifying and formally defining a set of inherent quality characteristics for LOD datasets based on existing standard quality models and frameworks. Subsequently a set of metrics is proposed which can effectively measure the inherent quality characteristics of LOD datasets. In the next section, in order to further clarify our notion of inherent quality in the context of LOD, we attend to different definitions of inherent quality characteristics, which have already been introduced in other existing quality models.

3 Inherent Quality Characteristics for LOD

The point of view that specifically considers the inherent quality characteristics of datasets is mainly concerned with real world representation and self-consistency. In other words, inherent quality is i) the ability of the dataset to effectively represent and reflect real world information in its data; and ii) the internal consistency of the data that are presented as a part of the same dataset. ISO/IEC 25012 defines the inherent viewpoint of data quality as "the degree to which quality characteristics of data have the intrinsic potential to satisfy stated and implied needs when data is used under specified conditions" [21]. In this point of view, data quality refers to the data itself, including data domain values and the possible relationships between data values and their metadata. According to ISO/IEC 25012, five characteristics are relevant from an inherent point of view, namely accuracy, completeness, consistency, credibility and timeliness.

Among these quality characteristics, accuracy, completeness and consistency are relevant to the inherent quality aspect of LOD, and are therefore considered in our study. Furthermore, accuracy is an important quality characteristic which has received the highest attention in the literature [21], [25], [27], [30] and it is also an important quality aspect of LOD. Therefore, to better investigate accuracy, we propose three different quality dimensions, namely semantic accuracy, syntactic accuracy and uniqueness. The other two quality dimensions, credibility and timeliness, are not included in our model, because they are not considered as inherent quality dimensions in the context of LOD. For example credibility is related to trustworthiness of the dataset, as well as other quality dimensions such as provenance, verifiability, believability, and licensing. Also, timeliness is an important aspect of LOD which is related to the dynamicity of a dataset and refers to the time during which the data is actually used. As result, it can be assessed at the time of usage and not at the time of publishing.

Therefore, based on the recommendation of ISO/IEC 25012 and in light of the above discussion, we adopt three of the characteristics proposed in this standard, namely accuracy, completeness and consistency. We further refine accuracy into three more fine grained characteristics; therefore, resulting in five main quality characteristics all of which represent some aspect of inherent quality of an open dataset. These inherent quality characteristics are namely semantic accuracy, syntactic accuracy, uniqueness, consistency and completeness and are shown in Figure 1. The definitions of these inherent quality characteristics are presented based on a systematic review of the literature and the ISO/IEC 25012 standard [21]. In the next sub-sections, we present an analysis of these inherent quality characteristics in the context of LOD.

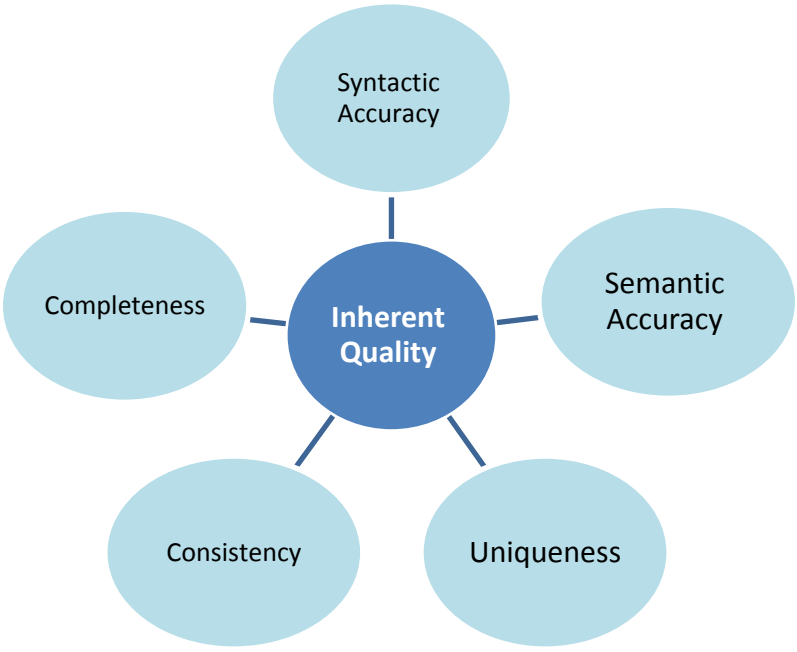

Figure 1: The proposed inherent quality characteristics of LOD

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762014000200006

3.1 Syntactic Accuracy

Data accuracy is highly related to other quality characteristics, such as data completeness, data consistency and data freshness, and there is no common agreement on its best definition; therefore, definitions are commonly overlapping. For example, the authors of [21] define accuracy from two main aspects namely, *syntactic accuracy* and *semantic accuracy*. Here, we study these two aspects of accuracy as separate quality characteristics. The authors in [27] state: "data is argued to be correct, in a syntactic way, if it satisfies syntactic rules and constraints imposed by the users." In [21], syntactic accuracy is defined as "the closeness of the data values to a set of values defined in a domain considered syntactically correct". Furthermore, syntactic accuracy can additionally be defined as the structural validity of a dataset, such as compliance with RDF/XML standard. In this study, we focus on the syntactic accuracy of entities as well as the appropriateness of the properties which are used for describing the entities, by measuring the erroneous or inaccurate annotations, classifications or representations.

3.2 Semantic Accuracy

Semantic accuracy relates to the correctness of a data value in comparison to the actual real world value. In [27], this aspect of accuracy is described as semantic correctness and "concerns the degree of correctness and validity of the data in comparison to the real world or with the reference data agreed to be correct." The ISO/IEC standard in [21] defines semantic accuracy as the "closeness of the data values to a set of values defined in a domain considered semantically correct", while [25] characterizes accuracy as "the percentage of objects without data errors such as misspellings, out-of-range values, etc." Another aspect of semantic accuracy is the precision of data. In [30], precision concerns the quantity of data to be stored and how precise this data must be. Generally, semantic accuracy means that every resource or entity described in a dataset should have a real world correspondence. In the context of LOD, semantic accuracy focuses on the correctness of data presented in a dataset, and consequently, in our study, resources referencing a wrong real world correspondent and entities with erroneous attribute values are examples of quality deficiencies related to semantic accuracy.

3.3 Uniqueness

Uniqueness is a quality characteristic that is rarely defined as an independent dimension in the literature. In [16] uniqueness is defined as the "degree to which data is free of redundancies, in breadth, depth and scope." Uniqueness in breadth is the degree to which ontology is free of redundancies regarding its represented classes and properties. Uniqueness in scope can be defined as the degree to which a knowledge base has multiple different instances to represent the same object. Uniqueness in depth represents the degree to which values of a property are unique; this is only suitable for properties that must assign each value of their domain only to one instance. In the context of LOD, uniqueness refers to the redundancy of concepts, either at the schema or the data level. At the schema level, it is similar to uniqueness in breadth defined in [16] and is related to the redundant classes and properties. At the data level, uniqueness means redundancies regarding both entities and values of properties. The latter has the exact meaning of uniqueness in depth defined in [16], but the former is different. Therefore, by uniqueness, we mean, uniqueness within the dataset as opposed to uniqueness in scope. In this way, uniqueness in the context of LOD is defined as the non-redundancy characteristic of the entities, classes, properties, and values of properties in a dataset.

3.4 Consistency

Consistency implies that "two or more values do not conflict with each other" [7]. In a generic way, it is defined as being free of conflicting information. Also, consistency is defined in [21] as "the degree to which data has attributes that are free from contradiction and are coherent with other data in a specific context of use. It can be either or both among data regarding one entity and across similar data for comparable entities". The consistency characteristic can be viewed from two perspectives, one being consistency of the same data values inside of a given dataset; and the other is consistency in the context of other datasets. Since we only intend to assess a dataset from an inherent point of view, we focus on the former and define it by the degree to which the format and the value of the data conform to the predefined schema. In addition, to better clarify the concept, we define consistency at two distinct levels, namely schema and data levels. Consistency at the schema level means that the schema of a dataset should be free of contradictions, while consistency at the data level focuses on the degree to which the format and the value of the data conform to the predefined schema of a given dataset.

3.5 Completeness

There are different definitions for completeness in the literature. For example, in [16] completeness refers to the degree to which all required information is presented in a particular dataset. The ISO/IEC standard in [21] defines completeness as "the degree to which subject data associated with an entity has values for all expected attributes and related entity instances in a specific context of use". In other words, completeness means that all the information

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762014000200006

required to describe a concept with a sufficient level of detail is presented. Although some other researchers consider the level of detail represented for description of data as precision [10], in our approach, the level of detail is defined as the amount of data available for the description of real world entities in a dataset, and is taken as a measure of completeness. In the context of LOD, completeness can be classified at two distinct levels, namely schema level and data level. At the schema level, completeness means that all of the required classes and properties should be represented while completeness at the data level refers to the missing values of properties with respect to the schema.

4 The Proposed Approach

In order to operationalize the measurement of the above five inherent quality characteristics, we define a set of metrics in this section. By employing the GQM (Goal Question Metric) approach, the process for metric definition starts by defining a set of goals, developing certain questions to characterize each of the goals, and is finalized by proposing metrics, which answer the posed questions.

4.1 Overview of GQM

GQM is a hierarchical model that starts with the definition of a set of goals that specify the purpose of measurement, the object to be measured, the issue of measurement, and the viewpoint from which the measurement is to take place. In this process, the goals are gradually refined into several questions that usually break down the issue into its major components. Each question is then refined into metrics. In some cases, one metric can be used to answer multiple questions. The resulting measurement model has three levels [2]:

- Conceptual level: In this level, specific goals are defined for an object of measurement for a variety of reasons, from different points of view relative to a particular environment.

- Operational level: A set of questions is developed to characterize the path to the achievement of a specific goal already defined in the conceptual level.

- Quantitative level: A set of metrics is defined to answer the questions that were defined in the operational level are addressed in a quantitative way.

Although the GQM approach was initially proposed in the software engineering field, it has since been widely applied in a variety of other domains. This approach has also become an accepted methodology for metric definition in the context of data quality assessment [31]. For example, [9] has applied traditional software metric definition techniques to measure the quality of data collections based on the GQM methodology. Likewise, the author of [33] has proposed a framework built on the GQM approach to propose certain quality metrics for managing the quality of data warehouses. In the following subsections, the process that we followed according to GQM for metric definition is presented.

4.2 Goal Definition

We define the following goal as the primary motivating objective for defining our metrics:

Goal: Assessment of the inherent quality of a dataset from the users' point of view in the context of LOD

This goal specifies a *purpose* (evaluate), the *object of measurement* (dataset), a *viewpoint* (user), an *issue* (inherent quality) and the *environment* (LOD) as required at the conceptual level of GQM. As mentioned earlier, we consider five characteristics for the inherent quality of LOD. Based on this classification, the main goal is decomposed into five sub-goals (SGs) as follow:

- Assessment of the *semantic accuracy* of a dataset from the users' perspective in context of LOD;

- Assessment of the *syntactic accuracy* of a dataset from the users' perspective in context of LOD;

- Assessment of the *uniqueness* of a dataset from the users' perspective in context of LOD;

- Assessment of the *consistency* of a dataset from the users' perspective in context of LOD;

- Assessment of the *completeness* of a dataset from the users' perspective in context of LOD.

In the following section, we address the above sub-goals by developing appropriate questions.

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762014000200006

4.3 Question Formulation

At the operational level, we define a set of questions to address the above sub-goals. We outline the formulated questions for each of the sub-goals in the following.

Semantic Accuracy:
Semantic accuracy means that every resource or entity described in a dataset of LOD should have a corresponding real world representation. To assess this aspect of accuracy, we need to show that every entity described in a dataset represents a real world object. Also, we should illustrate that all of the attributes used to describe the entities contain correct values. Based on these considerations the following questions are developed:

- Are the entities described with the correct values?
- Do the entities accurately represent the real world?

Syntactic Accuracy:
To assess syntactic accuracy, we need to demonstrate that a dataset dose suffer from syntactic errors. To do so, the following questions are developed:

- Is the syntax of the RDF documents valid?
- Are the resources described with the appropriate properties?

Uniqueness:
To assess the uniqueness characteristic of a given dataset, the degree of redundancy should be checked at both levels of the schema and data. At the schema level, uniqueness is related to redundant classes and properties, while at data level, it refers to the redundant entities and values of properties. Based on these two kinds of redundancy, the following questions are defined:

- What is the degree of redundancy in the context of classes?
- What is the degree of redundancy in the context of properties?
- Does a dataset contain multiple representations of the same entity?
- Does a dataset contain redundant values for properties?

Consistency:
As mentioned earlier, consistency can be viewed at both schema and data levels. At the schema level, consistency means that the schema of a dataset should be free from contradictions; while consistency at the data level refers to the inconsistency of available data entities and their properties values. Based on this definition, to assess the consistency of a dataset, the following two questions need to be answered:

- Is there any inconsistency in the schema of the dataset?
- What is the degree of conflict in the context of data values?

Completeness:
Similar to the consistency characteristic, completeness is defined at two levels of the schema and data. At the schema level it means that all of the required classes and properties should be represented, while completeness at the data level refers to the missing values of properties with respect to the schema. Therefore, the questions for completeness assessment of a given dataset are defined as follow:

- Have all the resources been described with adequate properties?
- Is all the required information for each entity presented?

In light of these questions, we propose metrics that can provide suitable answers to these questions.

Journal of Theoretical and Applied Electronic Commerce Research
ISSN 0718–1876 Electronic Version
VOL 9 / ISSUE 2 / MAY 2014 / 64-79

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762014000200006

4.4 Definition of Metrics

At the last step of GQM, a set of metrics should be defined in order to address the derived questions in a quantitative way. A data quality assessment metric or measure is a procedure for measuring an information quality characteristic [8]. Considering the fact that only few studies have been conducted which define the quality metrics for LOD [4], [18], [34] we need to define the required metrics from scratch and prior work cannot be reused for our purpose. We propose 20 metrics as measurement references for the inherent quality of open datasets. The main idea behind the design of these metrics has been comprehensiveness and simplicity. It should be noted that comparable to the metrics in other domains such as software engineering, the proposed metrics may not be comprehensive and future research might be required to complete this set by defining new metrics from other perspectives. This set of metrics is motivated and driven by the five sub-goals and the respective questions that were developed in the previous steps. The descriptions of the proposed metrics for semantic accuracy, syntactic accuracy, uniqueness, consistency and completeness are presented in Tables 1-5, respectively. The structure of Tables 1-5 represent the GQM format where the goals, questions and metrics are all included as they relate to each other.

Table 1: Proposed metrics for semantic accuracy

SG 1. Assessment of the semantic accuracy of a dataset from the users' perspective in the context of LOD		
Question	**Metric**	**Definition**
Are the entities described with the correct values?	*M 1.* Ratio of triples contain missing objects	The number of triples containing properties with missing values with respect to the properties defined in the schema divided by the total number of triples in a dataset subtracted from 1.
	M 2. Ratio of triples with out-of-range objects	The number of triples containing properties with outlier values with respect to the ranges of acceptable values defined in the schema divided by the total number of triples in a dataset subtracted from 1.
	M 3. Ratio of triples contain misspelling data value	The number of triples containing properties with misspelling values divided by the total number of triples in a dataset subtracted from 1.
Do entities accurately represent the real world?	*M 4.* Ratio of entities without correspondent in real-world	The number of entities without corresponding object in real-world divided by the total number of entities described in a dataset subtracted from 1.

Table 1 provides the definitions of the four metrics for semantic accuracy. As shown, the first three metrics are defined to answer the question about the attributes used to describe the entities, and the last metric is about the existence of the described entities in the real world.

Table 2: Proposed metrics for syntactic accuracy

SG 2. Assessment of the syntactic accuracy of a dataset from the users' perspective in the context of LOD		
Question	**Metric**	**Definition**
Is the syntax of the RDF documents valid?	*M 5.* Ratio of syntactically incorrect triples	The number of triples containing properties with syntactic error divided by the total number of triples in a dataset subtracted from 1.
Are the resources described with the appropriate properties?	*M 6.* Ratio of triples with improper assignments of data types to literals	The number of triples containing properties with improper assignment of data type to literal divided by total number of triples in a dataset subtracted from 1.
	M 7. Ratio of instances using undefined classes/properties	The number of instances using undefined classes/properties divided by total number of entities in a dataset subtracted from 1.
	M 8. Ratio of instances being Members of disjoint classes	The number of instances being members of disjoint classes divided by total number of instances in a dataset subtracted from 1.
	M 9. Ratio of triples containing improper usage of vocabularies	The number of triples which contain properties with improper usage of vocabularies divided by total number of triples in a dataset subtracted from 1.

Journal of Theoretical and Applied Electronic Commerce Research
ISSN 0718–1876 Electronic Version
VOL 9 / ISSUE 2 / MAY 2014 / 64-79

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762014000200006

Table 2 provides the definitions of the five metrics for syntactic accuracy. As shown, the first metric is defined to answer the question about the syntactic validity of the dataset, while the others are related to appropriateness of the properties used to describe the available entities.

Table 3: Proposed metrics for uniqueness

SG 3. Assessment of the uniqueness of a dataset from the users' perspective in the context of LOD		
Question	**Metric**	**Definition**
What is the degree of redundancy in the context of classes?	*M 10.* Ratio of redundant classes	The number of classes with the same properties divided by total number of classes defined in the schema of a dataset subtracted from 1.
What is the degree of redundancy in the context of properties?	*M 11.* Ratio of similar properties	The number of similar properties divided by total number of properties defined in the schema of a dataset subtracted from 1.
Does the dataset contain multiple representations for the same entity?	*M 12.* Ratio of redundant instances	The number of instances with different URI but the same properties divided by total number of instances in a dataset.
Does the dataset contain redundant values for the properties?	*M 13.* Ratio of functional properties with different values	The number of functional properties containing different values divided by total number of properties defined in the schema of a dataset subtracted from 1.

Table 3 shows the proposed metrics for the uniqueness characteristic. The goal of all these metrics is to check the degree of redundancy. The first two metrics are presented to answer the questions at the schema level by identifying redundant classes and properties, while the last two metrics refer to the redundant entities and values of properties at the data level.

Table 4: Proposed metrics for consistency

SG 4. Assessment of the consistency of a dataset from the users' perspective in the context of LOD		
Question	**Metric**	**Definition**
Is there any inconsistency in the schema of the dataset?	*M 14.* Membership of disjoint classes	The number of entities being members of disjoint classes.
	M 15. Invalid usage of inverse-functional properties	The number of triples containing invalid usage of inverse-functional properties.
	*M 16.*Ratio of triples using similar properties	The number of triples using similar properties divided by total number of triples in a dataset subtracted from 1.
	M 17. Heterogeneous data types	The number of heterogeneous data types in the schema.
What is the degree of conflict in the context of data value?	*M 18.* Inconsistent values of properties	The number of properties with inconsistent values.

Table 4 illustrates five metrics for the consistency characteristic. To propose these metrics, we have tried to focus on the consistency of a given dataset itself and cover all of the characteristics of an open dataset in this area. All of these metrics illustrate undesirable outcomes by enumerating the errors related to consistency.

Journal of Theoretical and Applied Electronic Commerce Research
ISSN 0718–1876 Electronic Version
VOL 9 / ISSUE 2 / MAY 2014 / 64-79

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762014000200006

Table 5: Proposed metrics for completeness

SG 5. Assessment of the completeness of a dataset from the users' perspective in the context of LOD		
Question	**Metric**	**Definition**
Have all of the resources been described with adequate number of properties?	*M 19.* Ratio of properties to class	The number of properties divided by the number of classes described in a dataset.
Is all the required information for each entity presented?	*M 20.* Missing properties per instance	Sum of the number of presented properties per instance divided by total number of instances in a dataset.

The proposed metrics for the completeness characteristic are shown in Table 5. At the schema level, completeness relates to the required classes and properties representing entities, and the corresponding metric computes the average number of properties used to describe classes, while completeness at the data level is evaluated by ratio of presented properties with respect to the schema of a dataset.

As shown in Tables 1-5, most of the metrics illustrate the desirable outcomes using the ratio scale. Simple ratio measures the ratio of the desired outcomes to total outcomes. A preferred way of computation is to calculate the number of the undesirable outcomes divided by that of the total outcomes subtracted from 1. This simple ratio adheres to the convention that 1 represents the most desirable and 0 the least desirable score [28]. The proposed metrics are classified into five groups according to the proposed quality characteristics introduced earlier. We have tried to cover as many of the inherent-related characteristics of an open dataset as possible.

5 Evaluation and Discussion

In this section, we support our claim for the suitability of the proposed metrics by considering both theoretical validation and empirical evaluation of our proposed metrics. Initially, the proposed metrics are validated from a measurement-theoretic perspective using the property-based measurement framework [11], [29], and subsequently, the comprehensiveness of the proposed metrics will be discussed. Furthermore, the quality characteristics depicted in Figure 1 are measured over eight real world datasets by applying our proposed metrics.

5.1 Theoretical Validation

Generally, any kind of measure is a homomorphism from an empirical relational system to a numerical relational system [13], [14]; therefore, it is imperative that measures be theoretically analyzed within the framework of measurement theory. There are two main groups of frameworks for the theoretical validation of metrics in the literature. The first group consists of frameworks directly based on measurement theory principles [29]; while the second group expresses the desirable properties of the numerical relational system that need to be satisfied by the metrics [11]. In the following, we have examined the properties of our metrics according to the framework proposed by Briand et al, which is one of the most well-known frameworks in the latter group. This property-based measurement framework [11] provides five types of metrics including *size*, *length*, *complexity*, *coupling* and *cohesion* and offers a set of desirable properties for each of these metric types.

Within the property-based measurement framework, size is recognized as being an important measurement concept and is used when metrics are defined based on numeration. One of the well-known size metrics is LOC (Lines of Code). Length measures can be defined at the system or module level based on the length concept. A typical example is the depth of a hierarchy. Our proposed metrics cannot be defined as length measures because all of the metrics are applied to RDF file, not RDF graph or any other hierarchical structure. The last two metric types are cohesion and coupling which are only relevant for modular systems and they are not applicable to our domain, because the proposed metrics are defined to be applied on individual datasets regardless of their linkage to the other datasets of LOD. Thus, based on the definitions of the proposed metrics presented in Tables 1-5, all of the metrics are of the complexity type. In [11], five desirable properties are provided for complexity metric type, namely, non-negativity, null value, symmetry, additivity and monotonicity. Table 6 analyzes these properties for our proposed metrics. As seen, all of the metrics respect non-negativity, null value, symmetry and monotonicity, but not additivity. We believe that although the defined complexity metrics do not exactly satisfy additivity, but their satisfaction of the monotonicity properties makes them acceptable for our purpose. Because additivity is a special case of monotonicity, i.e., mono-tonicity is less strict than additivity. Hence, all of the proposed metrics respect four out of the five desirable properties defined by the property-based measurement framework [11] to form valid metric space.

Journal of Theoretical and Applied Electronic Commerce Research
ISSN 0718–1876 Electronic Version
VOL 9 / ISSUE 2 / MAY 2014 / 64-79

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762014000200006

Table 6: Theoretical analysis of the proposed metrics

Metrics (shown in Tables 1-5)	Desirable properties				Measure Scale	Metric Type	Method of Measurement
	Non-negativity	Null value	Symmetry	Monotonicity	Nominal/ Ordinal/ Interval/ Ratio	Subjective/ Objective	Automated/ Expert/ Validator
M1	+	+	+	+	R	O	A
M2	+	+	+	+	R	O	A
M3	+	+	+	+	R	O	A
M4	+	+	+	+	R	S	E
M5	+	+	+	+	R	O	V
M6	+	+	+	+	R	O	A
M7	+	+	+	+	R	O	A
M8	+	+	+	+	R	O	A
M9	+	+	+	+	R	S	E
M10	+	+	+	+	R	O	A
M11	+	+	+	+	R	O	A
M12	+	+	+	+	R	O	A
M13	+	+	+	+	R	O	A
M14	+	+	+	+	R	O	A
M15	+	+	+	+	R	O	A
M16	+	+	+	+	R	O	A
M17	+	+	+	+	R	O	A
M18	+	+	+	+	R	O	A
M19	+	+	+	+	R	O	A
M20	+	+	+	+	R	O	A

Table 6 also depicts additional features of the proposed metrics such as scale, type and method of measurement. As mentioned earlier, all of the metrics are measured by computing the desirable outcomes using either ratio or the number of deficiencies in a dataset. Therefore, the important consequence of satisfying these properties is that all of our proposed metrics are in ratio scale [13], which means that they are theoretically valid metrics. From the type perspective, most of the metrics are Objective (O) and can be systematically measured. As a result the methods of measurement for these metrics are developing code or using validators. Only two metrics are Subjective (S) and should be measured by receiving experts' subjective perception using methods such as questionnaire.

5.2 Empirical Evaluation

To further study the behavior of the proposed metrics, it is necessary to place them under empirical evaluation and show their applicability in practice. In this section, we report the results of our observations with regard to the calculation of the proposed metrics for several real world datasets. We further report on our statistical analysis of the correlation between the proposed metrics.

We have selected eight datasets and made sure that these datasets: *i)* were from across a variety of domains; and *ii)* were of different sizes in terms of the number of triples in the datasets as shown in Table 7. The datasets were selected from the EU FP6 Networked Ontology (NeOn) project (Site 3) with the restriction that the language of the datasets needed to be English.

The main purpose of the study was to identify the relationships between the proposed metrics and inherent quality characteristics of datasets to be released to the LOD cloud. For this purpose, we calculated the values of the metrics for each of the eight datasets in order to observe the trends of the metrics over datasets of various domains and sizes. The calculation of the values for each metric was accomplished automatically. We have implemented an automated tool that is able to automatically compute the metric values for any given input dataset. The datasets and the code are available online at [22]. Tables 8 and 9 present all of the collected values of the proposed metrics for each of the datasets.

Journal of Theoretical and Applied Electronic Commerce Research
ISSN 0718–1876 Electronic Version
VOL 9 / ISSUE 2 / MAY 2014 / 64-79

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762014000200006

Table 7: The details of the datasets used in our experiments

Datasets	No. of triples	No. of instances	No. of classes	No. of properties
FAO Water Areas	5,365	293	7	19
Water Economic Zones	25,959	693	22	127
Large Marine Ecosystems	6,006	358	9	31
Geopolitical Entities	22,725	312	11	101
ISSCAAP Species Classification	368,619	23,856	22	93
Species Taxonomic Classification	318,153	11,738	5	26
Commodities	28,210	1,394	6	19
Vessels	2,118	120	6	22

Table 8: Observations for Metrics 1 – 10

Metrics / Datasets	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10
FAO Water Areas	0.67	0.84	0.84	N/A	1.00	1.00	0.50	0.22	N/A	0.71
Water Economic Zones	0.26	0.81	0.99	N/A	1.00	1.00	0.63	0.52	N/A	0.23
Large Marine Ecosystems	0.44	0.78	0.85	N/A	1.00	1.00	0.53	0.36	N/A	0.88
Geopolitical Entities	0.59	0.78	0.99	N/A	1.00	1.00	1.00	0.00	N/A	0.45
ISSCAAP Species Classification	0.15	0.12	0.95	N/A	1.00	0.58	0.94	0.95	N/A	0.95
Species Taxonomic Classification	0.62	0.04	0.95	N/A	1.00	0.52	1.00	0.92	N/A	1.00
Commodities	0.70	0.7	0.85	N/A	1.00	1.00	0.46	0.90	N/A	0.00
Vessels	0.66	0.11	0.83	N/A	1.00	1.00	0.50	0.23	N/A	0.50
Mean	0.51	0.52	0.91	-	1.00	0.89	0.69	0.51	-	0.59
STDEV	0.21	0.36	0.07	-	0.00	0.21	0. 24	0.37	-	0.36

As shown in Table 8, the values for Metrics 4 and 9 are not reported due to their subjective nature. Both of these metrics are Subjective (S) and require the intervention of an expert for their measurement. Metric 4 identifies the number of entities that do not have correspondences in the real world, and Metric 9 checks the vocabularies used in a dataset and investigates whether a given dataset has used the proper vocabularies to describe the data or not. It is clear that both of these metrics are very subjective and it is hard, if not impossible, to asses them automatically. For this reason, we were not able to compute the values of these metrics using our automated metric computation tool. The best way to compute these two metrics is to receive experts' subjective perception using methods such as questionnaires. A major limitation of such an approach would be the effect of the subjective nature of expert opinions, which can be in some cases inconsistent. We are planning to investigate the best approach for the collection of these two metrics in a semi-automated way in the future.

Journal of Theoretical and Applied Electronic Commerce Research
ISSN 0718–1876 Electronic Version
VOL 9 / ISSUE 2 / MAY 2014 / 64-79

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762014000200006

Table 9: Observations for Metrics 11 – 20

Metrics / Datasets	M11	M12	M13	M14	M15	M16	M17	M18	M19	M20
FAO Water Areas	1.00	1.00	1.00	228.00	1.00	1.00	5.00	1,047	2.71	0.67
Water Economic Zones	0.95	1.00	1.00	331.00	1.00	0.95	11.00	4,875	5.77	0.24
Large Marine Ecosystems	1.00	1.00	1.00	228.00	1.00	1.00	5.00	1,133	3.44	0.44
Geopolitical Entities	0.94	1.00	1.00	312.00	1.00	0.94	7.00	4,340	9.18	0.59
ISSCAAP Species Classification	0.87	1.00	1.00	1,127.00	1.00	0.89	7.00	115,410	4.22	0.13
Species Taxonomic Classification	0.88	1.00	1.00	983.00	1.00	0.90	2.00	105,011	5.20	0.63
Commodities	0.84	1.00	1.00	144.00	1.00	0.85	2.00	10,399	3.16	0.70
Vessels	0.91	1.00	1.00	93.00	1.00	0.91	1.00	360	3.66	0.66
Mean	0.92	1.00	1.00	430.75	1.00	0.93	5.00	30,321.88	4.67	0.51
STDEV	0.06	0.00	0.00	395.02	0.00	0.05	3.34	49,489.49	2.09	0.22

In light of the values of the metrics as reported in Tables 8 and 9, it is clear that some of the proposed metrics have the same value regardless of the dataset. For this reason, we conclude that these metrics are not discriminative metrics and therefore they would not be suitable metrics for determining the inherent quality of a dataset. These metrics are M5, M8, M13 and M15. Given the fact that the role of a metric is to measure an aspect of a dataset that would distinguish it from other datasets as much as possible and in light of the fact that these four metrics have a standard deviation of zero over all of the eight experimented datasets, we recommend that these metrics are not used when making measurements over LOD datasets. We provide further details as follows:

- Metric 5 (M5) is defined to collect the number of triples containing properties with syntactic error and it can be done using online validators. These validators get an RDF/XML document as input and check if the document is syntactically valid. The results of performing validation on the NeOn datasets illustrated that all of the datasets are syntactically correct .Thus, the values of this metric are 1 for all of them. It means that selected datasets have no syntax errors. In most cases, datasets are built using tool support, which ensure that the datasets are syntactically correct.

- Metric 12 (M12) refers to the unique instances and is measured by the ratio of the number of similar instances (instances with different URIs, but the same properties) to the total number of instances in a dataset. It intends to address cases when two datasets are merged together and sometime an entity exists in both of them. Similar to M5, tool supports usually identify and resolve such duplicates; this is the reason why the value for this metric is 1 in all of the datasets.

- Metric 13 (M13) refers to the invalid usage of functional properties by measuring the ratio of the functional properties containing different values in a dataset. The metric's value 1 for all of the datasets indicates that there is not any functional property with different values in all of the datasets.

- Metric 15 (M15) is similar to Metric 13, with a slight difference that it refers to the invalid usage of inverse functional properties and the reason for number of the triples functional properties containing different values in a dataset

Based on this observation, our recommendation is to eliminate these four metrics from the set of proposed metrics. The other important aspect of the metrics that we would like to investigate is the inter-metric correlation, which shows how much the proposed metrics are related to each other. Ideally, we would like our metrics to be independent from each other in a way that each of them represents a distinct aspect of the inherent quality of datasets. In order to measure inter-metric correlation, we perform a non-parametric test namely the Spearman's Rho correlation. Spearman's correlation is used to identify the relationship between the defined metrics. In other word this test is used to see whether some of the metrics capture similar aspects of the datasets and whether they are overlapping or not.

Journal of Theoretical and Applied Electronic Commerce Research
ISSN 0718–1876 Electronic Version
VOL 9 / ISSUE 2 / MAY 2014 / 64-79

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762014000200006

Table 10: Inter-metric correlation for the proposed metrics

Rho	**M1**	**M2**	**M3**	**M6**	**M7**	**M8**	**M10**	**M11**	**M14**	**M16**	**M17**	**M18**	**M19**	***M20***
M1		-0.04	-0.62	0.31	-0.63	-0.31	-0.33	-0.14	-0.72	-0.14	-0.73	-0.38	-0.60	1.00
p value		+	+	+	+	+	+	+	-	+	-	+	+	-
M2			0.12	0.70	-0.22	-0.54	-0.42	0.72	-0.10	0.72	0.60	-0.37	-0.18	-0.04
p value			+	-	+	+	+	-	+	-	+	+	+	+
M3				-0.25	0.76	0.18	-0.09	-0.08	0.74	-0.08	0.76	0.59	0.84	-0.62
p value				+	-	+	+	+	-	+	-	+	-	+
M6					-0.59	-0.73	-0.76	0.49	-0.74	0.49	0.06	-0.73	-0.26	0.31
p value					+	-	-	+	-	+	+	-	+	+
M7						0.13	0.48	-0.01	0.81	-0.01	0.46	0.46	0.83	-0.63
p value						+	+	+	-	+	+	+	-	+
M8							0.36	-0.65	0.53	-0.65	-0.02	0.83	0.02	-0.31
p value							+	-	+	-	+	-	+	+
M10								0.06	0.51	0.06	-0.12	0.24	0.00	-0.33
p value								+	+	+	+	+	+	+
M11									-0.13	**1.00**	0.32	-0.66	-0.10	-0.14
p value									+	-	+	-	+	+
M14										-0.13	0.62	0.75	0.57	-0.72
p value										+	+	-	+	-
M16											0.32	**-0.66**	-0.10	-0.14
p value											+	-	+	+
M17												0.29	0.50	-0.73
p value												+	+	-
M18													0.33	-0.38
p value													+	+
M19														-0.60
p value														+

(+ means *p-value >0.05* and – means *p-value ≤ 0.05*)

The results of this study are shown in Table 10. There are two rows dedicated to each metric in this table; the values on the upper rows show the degree of correlation (rho value), while the symbols on the lower rows depict the significance of the correlation (p value). The meanings of symbols are shown below the table. As usual, a significance level of $p < 0.05$ (95% confidence level) is used to accept the results of the correlation. According to Spearman's correlation, a correlation with a significance p value ≤ 0.05 can be considered to be significant. Therefore, in our work such correlations are considered to be meaningful and are depicted with the symbol of -, while the symbol of + indicates that there are no meaningful correlations. Amongst them, those values with the higher degree of correlation are regarded as the most significant correlation and are underlined.

The results illustrate that some of the metrics are in fact correlated and are perhaps overlapping in terms of the concepts that they are representing. For instance, M11 and M16 are highly correlated, which is an indication that these two metrics are similar for a dataset. In other words, a dataset with higher number of similar properties is highly likely to have more triples using these properties. On the other hand, a meaningful negative correlation can be seen between M16 and M18. This means that using similar properties have an inverse relation with the inconsistency of data values in a dataset. The correlations values found between the metrics show that most of the proposed metrics are in fact not correlated and therefore stand to represent independent aspects of a dataset. Therefore, the metrics would be quite suitable for a comprehensive evaluation of the inherent quality characteristics of a dataset from different aspects due to the fact that almost 70% of the values presented in the Table 10 are depicted with +, which means that their corresponding p-values are more than 0.05.

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762014000200006

6 Conclusion

In this paper, a set of measurement-theoretic metrics have been proposed for evaluating the inherent quality characteristics of a dataset. These quality characteristics are namely semantic accuracy, syntactic accuracy, uniqueness, completeness and consistency. In the first part of our work, we have reviewed studies that have been reported on various aspects of data quality in the two main subgroups of information quality models and data quality models in the context of LOD. We have also discussed quality characteristics, which depict our notion of inherent quality in the context of LOD. Subsequently, we have discussed how concrete valid metrics can be developed for open datasets based on the GQM method. Finally, the suitability of the proposed metrics has been discussed through both theoretical validation and empirical evaluation.

To evaluate the quality of a dataset, it is crucial to consider various stakeholders' needs. The final objective of our work is both to help data providers in evaluating their data as well as enable users to filter out poor quality data. Besides the publisher and user roles, third party applications or mashups are the other important stakeholders. If the datasets on the LOD are incorrect or incomplete, this may lead to significant issues at the application level. Therefore, the metrics are naturally defined according to this goal by considering the various viewpoints of all stockholders. Finally, with respect to the variety of domains in LOD, generality is a crucial principle for the methods of quality assessment. Since the proposed metrics can be applied for quality evaluation of the datasets in any domain, domain independence of the defined metrics is another advantage of the solution offered in the present work.

We are currently focusing on the extension of our work in two main directions: *i)* we are working on finding the best approach to gather values for the metrics that have a subjective nature and cannot be automatically computed from the structure of the datasets. Metrics 4 and 9 are such metrics and we plan to develop suitable semi-automated methods for capturing their value for different datasets; and also *ii)* we are considering to develop statistical models that would take the values of the metrics proposed in this paper for each dataset into account and predict the possible inherent quality of the dataset once it is integrated into the LOD. We have undertaken similar studies for building predictive models of quality from structural metrics in our prior research [1] .

Acknowledgments

The authors would like to gratefully acknowledge the support from Ryerson University's Laboratory for Systems, Software and Semantics (LS^3) in particular Javad Ghatta, Mahdi Bashari and Jelena Jovanovic.

Websites list

Site 1: Vapour online validator
http://validator.linkeddata.org/vapour

Site 2: Jena eyeball project
https://jena.apache.org/documentation/tools/eyeball-getting-started

Site 3: Networked Ontology (NeOn) project
http://www.neon-project.org

References

[1] E. Bagheri and D. Gasevic, Assessing the maintainability of software product line feature models using structural metrics, Software Quality Journal, vol. 19, pp. 579-612, 2011.
[2] V. R. Basili, G. Caldiera, and H. D. Rombach, The goal question metric approach, in Encyclopedia of software engineering, ed: John Wiley & Sons, 1994, pp. 528-532.
[3] C. Batini and M. Scannapieca, Data quality: concepts, methodologies and techniques: Springer, 2006.
[4] C. Batini, C. Cappiello, C. Francalanci, and A. Maurino, Methodologies for data quality assessment and improvement, ACM Computing Surveys (CSUR), vol. 41, p. 16, 2009.
[5] F. Bauer and M. Kaltenböck, Linked Open Data: The Essentials, Edition mono/monochrom, Vienna, 2011.
[6] B. Behkamal, M. Kahani, S. Paydar, M. Dadkhah, and E. Sekhavaty, Publishing Persian linked data; challenges and lessons learned, in Telecommunications (IST), 2010 5th International Symposium on, 2010, pp. 732-737.
[7] C. Bizer, Quality Driven Information Filtering: In the Context of Web Based Information Systems: VDM Publishing, 2007.
[8] C. Bizer and R. Cyganiak, Quality-driven information filtering using the WIQA policy framework, Web Semantics: Science, Services and Agents on the World Wide Web, vol. 7, pp. 1-10, 2009.

This paper is available online at
www.jtaer.com
DOI: 10.4067/S0718-18762014000200006

[9] M. Bobrowski, M. Marré, and D. Yankelevich, A software engineering view of data quality, Intl. Software Quality Week Europe (QWE'98), 1998.
[10] M. Bobrowski , M. Marré, and D. Yankelevich, A Homogeneous Framework to Measure Data Quality, presented at the International Conference on Information Quality (IQ), Cambridge, MA, 1999.
[11] L. C. Briand, S. Morasca, and V. R. Basili, Property-based software engineering measurement, Software Engineering, IEEE Transactions on, vol. 22, pp. 68-86, 1996.
[12] S. Brüggemann and F. Grüning, Using ontologies providing domain knowledge for data quality management, in Networked Knowledge-Networked Media, ed: Springer, 2009, pp. 187-203.
[13] N. E. Fenton and S. L. Pfleeger, Software metrics: a rigorous and practical approach: PWS Publishing Co., 1998.
[14] L. Finkelstein, Widely, strongly and weakly defined measurement, Measurement, vol. 34, pp. 39-48, 2003.
[15] A. Flemming and O. Hartig. (April, 2014) URI Debugger. Linked Data Informatik. [Online]. Available: http://linkeddata.informatik.hu-berlin.de/uridbg
[16] C. Fürber and M. Hepp, Using semantic web resources for data quality management, in Knowledge Engineering and Management by the Masses, ed: Springer, 2010, pp. 211-225.
[17] C. Fürber and M. Hepp, SWIQA–A Semantic Web information quality assessment framework, 2011.
[18] O. Hartig, Trustworthiness of data on the web, in Proceedings of the STI Berlin & CSW PhD Workshop, 2008.
[19] O. Hartig and J. Zhao, Using Web Data Provenance for Quality Assessment, SWPM, vol. 526, 2009.
[20] A. Hogan, A. Harth, A. Passant, S. Decker, and A. Polleres, Weaving the pedantic web, 2010.
[21] ISO, "ISO/IEC 25012- Software engineering - Software product Quality Requirements and Evaluation (SQuaRE)," in Data quality model, ed, 2008.
[22] Z. Jeremic (July 2013). The code of metrics calculation tool (1.0 ed.). Atlassian Bitbucket. [Online]. Available: https://bitbucket.org/behkamal/new-metrics-codes/src
[23] Y. W. Lee, D. M. Strong, B. K. Kahn, and R. Y. Wang, AIMQ: a methodology for information quality assessment, Information & management, vol. 40, pp. 133-146, 2002.
[24] Y. Lei, A. Nikolov, V. Uren, and E. Motta, Detecting Quality Problems in Semantic Metadata without the Presence of a Gold Standard, in 5th International EON Workshop at International Semantic Web Conference (ISWC'07), Busan, Korea, 2007, pp. 51-60.
[25] F. Naumann and C. Rolker, Assessment methods for information quality criteria, in 5'th Conference on Information Quality Cambridge, MA, 2000, pp. 148-162.
[26] F. Naumann, U. Leser, and J. C. Freytag, Quality-driven integration of heterogeneous information systems, in 25th International Conference on Very Large Data Bases (VLDB'99), Edinburgh, Scotland, UK, 1999, pp. 447-458.
[27] V. Peralta, Data freshness and data accuracy: A state of the art, Instituto de Computacion, Facultad de Ingenieria, Universidad de la Republica, Uruguay, Tech. Rep. TR0613, 2006.
[28] L. L. Pipino, Y. W. Lee, and R. Y. Wang, Data quality assessment, Communications of the ACM, vol. 45, pp. 211-218, 2002.
[29] G. Poels and G. Dedene, Distance-based software measurement: necessary and sufficient properties for software measures, Information and Software Technology, vol. 42, pp. 35-46, 2000.
[30] T. C. Redman and A. Blanton, Data quality for the information age: Artech House, Inc., 1997.
[31] S. A. Sarcia, Is GQM+ Strategies really applicable as is to non-software development domains?, in Proceedings of the 2010 ACM-IEEE International Symposium on Empirical Software Engineering and Measurement, 2010, p. 45.
[32] K. Tolle. (April, 2014) The validating RDF parser (VRP) (3.0 ed.). Forth Institute of Computer Science. [Online]. Available: http://139.91.183.30:9090/RDF/VRP
[33] P. Vassiliadis, Data warehouse modeling and quality issues, Ph.D. dissertation, Department of Electrical and Computer Engineering, National Technical University of Athens Athens, GREECE, 2000.
[34] A. Zaveri, D. Kontokostas, M. A. Sherif, L. Bühmann, M. Morsey, S. Auer, et al., User-driven quality evaluation of dbpedia, in Proceedings of the 9th International Conference on Semantic Systems, 2013, pp. 97-104.