User Profiles in Information Web Portals

Carmen Moraga, Mª
Ángeles Moraga
Alarcos Research Group
University of Castilla-La
Mancha,
Ciudad-Real, Spain
Carmen.Moraga@alu.uclm.es
MariaAngeles.Moraga@uclm.es

Angélica Caro
Department of Computer
Science and Information
Technologies, University of
Bio Bio,
Chillán, Chile
mcaro@ubiobio.cl

Coral Calero
Alarcos Research Group
University of Castilla-La
Mancha,
Ciudad-Real, Spain
Coral.Calero@uclm.es

Rodrigo Romo Muñoz
Department of Business
Management, University of
Bio Bio,
Chillán, Chile
rromo@ubiobio.cl

Abstract—**The number of Web portals is increasing daily. These Web portals can be grouped into different types according to their purpose. One of these types is 'Information Web Portals' in which data quality is particularly important to users. This paper uses a survey to study the relevance that users place on a series of data quality characteristics in this type of Web portal. To do this, we determined various user profiles based on demographic aspects (gender, age range, level of studies and type of organization). We also analysed whether each profile prefers some characteristics to others. The results obtained will allow designers and developers to know which data quality characteristics they should place most emphasis on depending on the users of a particular Web portal.**

Keywords-Data Quality; Web portal; Statistical Method.

I. INTRODUCTION

The Internet has become a powerful tool for communication purposes, both for the exchange of information and ideas and for learning and gaining knowledge, and it is even used for participation in local, national and international networking. The Internet has also come to be used in all aspects of our life in recent years [1]. One means to access information on the Internet is through Web portals. A Web portal is an entry point to the Internet [2-4]. Web portals select, organize, integrate distributed contents and enable viewers to access organizational services via the Internet [5-7]. Web portals play an increasingly specialized role in the online world [3] and they also allow the creation of a work and/or business environment in which users can navigate in a simple manner to rapidly obtain the information that they need [8], thus facilitating access to data on the Internet.

We have found several classifications of Web portal (i.e [9, 10]) depending on the type of purpose. In our work, we have divided Web portals into the following groups according to the principal type of activity that users wish to carry out:

- 'The Search for and Reading of Information': defined as those portals that the user uses solely to obtain information (e.g., a TV channel Portal to discover what programs are being shown, a cinema Portal to see what films are being shown, a newspaper Portal,

etc.) This type of portal is, therefore, merely informative.
- 'Commercial Interaction': determined by the fact that it is used to carry out some kind of on-line transaction, such as buying train or airline tickets, making downloads of a legal nature, transferring money, making payments, etc. This type of portal is, therefore, of a transactional nature.
- 'Interaction with other People': the important aspect here is the ability to relate to or get in contact with other people, known or otherwise. For example, social networks. This type of portal is, therefore, of the data-exchange type.

Although the same users may access each type of Web portal, their preferences may be different according to, amongst other things, their demographic aspects, for example, with regard to gender [11], [12]. Men are, according to [13], more analytical and therefore more objective, whilst women are more subjective and intuitive. The level of the user's studies also influences Web portal use and, as the author states in [14], there is a significant difference between the Internet addiction scores of students and other professional groups.

Furthermore, Data Quality (DQ), which is often defined as the ability of a collection of data to meet user requirements [15, 16], is increasingly more important to Internet users [15, 17, 18] . This importance resides in the fact that users can use the data obtained to carry out everyday tasks and to make decisions both in their jobs and in their personal lives. With this, the area of Web portal data quality has consequently begun to emerge [19]. The Web portal owners are aware that DQ is important to increase user reliability, since users can clearly see its usefulness. Thus, when the degree of satisfaction increases, the number of customers that access the portal also increases.

All of the above led us to consider that it would be interesting to establish whether different user profiles exist as regards preferences towards the various characteristics of DQ in a Web portal. This was done by studying the following demographic aspects: the gender, age range, level of studies and type of organization to which Web portals users are linked.

In this paper, we particularly focus on the analysis of DQ characteristics for Web portals of the 'Search for and Reading

of Information' type (which from here on will be referred to as 'Information Web Portals'). The DQ reference model used was SPDQM (SQuaRE-Aligned Portal Data Quality Model), a DQ model for Web portals which provides 42 DQ characteristics. These DQ characteristics are distributed in 4 categories: Intrinsic, Operational, Contextual and Representational. [20]. We shall use the set of DQ characteristics in the Contextual category, since it is the only category in which the importance that users place on certain DQ characteristics with regard to others may vary according to the type of Web portal.

Our study will allow us to determine, first, if all the DQ characteristics are important for users, and second, which are most relevant in comparison to the others according to the different user profiles established.

The results obtained will allow designers and developers to discover the most relevant DQ characteristics according to the aforementioned user profiles, in order to reinforce these DQ characteristics in the Web portal and thus satisfy their users.

The remainder of this paper is organized as follows: in Section 2, the DQ characteristics in the Contextual category are defined. Section 3 describes the data-collection method used. The various user profiles and their preferences as regards the DQ characteristics in the 'Information Web Portals' is determined in Section 4. In Section 5, guidelines for designers and developers are indicated. Finally, Section 6 presents our conclusions and future work.

II.   THE CONTEXTUAL DQ IN WEB PORTALS

As already mentioned, in this work we shall focus on 'Information Web Portals' to determine the relevance of the DQ characteristics in the SPDQM Contextual category.

In SPDQM, the Contextual category highlights the requirement which states that data quality must be considered within the context of the task at hand. A certain type of context must therefore be considered in the Contextual category in order to establish the data quality in this environment, in our case, the 'Information Web Portals'.

This category contains 10 DQ characteristics and 6 DQ sub-characteristics (see Fig. 1). Further information on SPDQM can be found in [20].

Moreover, the importance that is placed on certain DQ characteristics in the Contextual category with regard to others may be different according to the various user profiles.
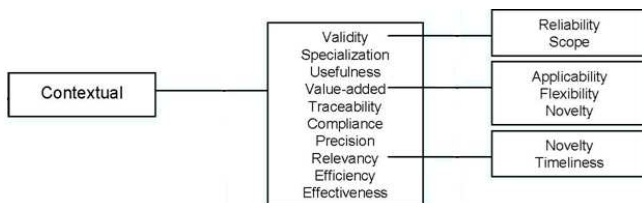

Figure 1. DQ characteristics in the contextual category

III.   DATA COLLECTION SURVEY

In order to determine the user profiles and the DQ priorities for the different profiles of 'Information Web portal' users, we decided to carry out an unsupervised survey, in which questions related to this type of Web portals appeared. We used "the principles of survey research" proposed in [21, 22].

The questionnaire in this survey was formed of a total of 21 questions, 4 of which were general, related to demographic aspects (Table I), 16 of which were related to DQ characteristics in the Contextual category and 1 question concerning the definition of the term 'Contextual'. The questions concerning the DQ characteristics had to be easy to understand. Pre-test questionnaire was therefore first carried out with users who were experienced in the use of Web portals and whose feedback allowed us to modify the initial questions and obtain a definitive set of understandable questions for all types of Web portal users (Table II). In the questionnaire, only one response was possible because we used closed questions.

The questions were answered using an 11-point Likert-type interval scale, ranging from strongly disagree (0) to strongly agree (10).

The final questionnaire was distributed to a heterogeneous group of 200 Web portal users from Europe and Latin America, by e-mail or in printed format. The questionnaires were collected in the same manner, and 192 of them were returned, signifying that a response rate of 96% was obtained. However, 4 surveys had to be discarded because they were incomplete. We were therefore left with 188 surveys that could be used, thus obtaining a response rate of 94%.

Once the data obtained with the questionnaires had been collected, it was necessary to carry out a statistical analysis to investigate the results [23] .

The starting point for this was the calculation of the Cronbach's alpha to estimate the reliability of the results. A value of 0.942 was obtained as a result of this, which indicated that the results had good internal consistence. The information was therefore reliable.

An analysis of the sample obtained is shown as follows.

IV.   DETERMINING PREFERENCES FOR DQ CHARACTERISTICS ACCORDING TO THE USER PROFILE

In this section we shall determine whether all the DQ characteristics identified in the Contextual category in our SPDQM model are considered to be important by the users of this type of Web portals, and we shall also identify whether any of these DQ characteristics are more important than others depending on the various user profiles.

This will be done by carrying out a statistical analysis of the results obtained from the survey with the use of an SPSS statistical analysis tool and on the basis of the following steps:

TABLE I.   QUESTIONS CONCERNING DEMOGRAPHIC ASPECTS

Gender: Male/Female
Level of studies COMPLETED: High School / Vocational Training/ University/ Post Graduate.
Type of organization with which you are linked (for study or work purposes). If there are various, please place them in the order in which most time is dedicated to them, from greatest to least: Education / Service Sector / Industrial - Commercial – Financial / Other (Please state which).
Age range: Under 25 / Between 25 and 35 / Between 35 and 45 / Between 45 and 55 / Over 55.

TABLE II.   QUESTIONS CONCERNING THE CONTEXTUAL DQ CATEGORY

1.- The data should be sufficiently detailed to facilitate the task at hand.
2.- The data obtained from a Web portal should be true and reliable (believable).
3.- In general, the data in Web portals should be understandable for you to consider them valid. (This characteristic is related to those shown in questions 1 and 2).
4.- It should be possible to verify the data, and it would be appropriate to know their author and/or their source, and to be able to obtain a record of any modifications made to them.
5.- The data should be defined in accordance with a regulation, i.e., they should comply with pre-established standards (by, for example, showing dates, prices, etc.) thus avoiding situations in which doubts or different interpretations emerge.
6.- The data provided by a Web portal should contain the appropriate and specific information for the use to which they will be put.
7.- The data should adapt to user needs (e.g., they should be integrated into other applications or presented in different formats).
8.- The data should be useful and specially oriented towards the user community that will utilize them.
9.- The data should be innovative, thus allowing those who use them to obtain new knowledge.
10.- In general, the data should be suitable to allow the user to obtain advantages thanks to their use. (This characteristic is related to those shown in questions 7, 8 and 9).
11.- The data should be available in the shortest possible amount of time.
12.- In general, the data should be applicable and innovative, and should be available in a reasonable amount of time. (This characteristic is related to those shown in questions 9 and 11).
13.- The data offered by portals should be useful for their users, and should satisfy their needs.
14.- It should be possible to obtain the data by using the appropriate quantities and types of resources (by, for example, using the smallest possible number of links to locate the data desired).
15.- The data obtained from Web portals should be exact and concise, thus helping you to find relevant results.
16.- The data in Web portals should provide the information that users are seeking.
17.- The data provided by Web portals should have a level of quality that accords with the specific use to which you wish to put them, i.e., in the context of the specific area in which you wish to work with them.

A. Study of DQ Characteristics: Descriptive Statistical

A descriptive statistical analysis was used to show the minimum, maximum and mean values of all the DQ characteristics (including a question concerning the definition of the term 'Contextual') with the objective of verifying whether all these DQ characteristics are important in 'Information Web Portals'. 11 characteristics with mean values over 8 were obtained, while the others have values of between 7.57 and 8, all of them being above 7.5. The mean values given by users are thus high (nearer to the maximum value of 10 than to the intermediate value of 5). This signifies that all the DQ characteristics in the Contextual category are, in effect, important to the users of 'Information Web Portals'.

B. Creation of groups of DQ characteristics: Factor analysis

This sub-section was carried out with the use of a factorial analysis which allowed us to determine homogeneous groups of DQ characteristics (denominated as factors). These factors permitted a summary to be made of the relationship between the DQ characteristics considered. Each factor is independent of the others.

The results reveal the existence of three factors (Table III). Factor 1, Factor 2 and Factor 3 represent 53.48%, 7.47% and 5.83% of the total variance, respectively. These factors, when combined, can therefore explain 66.77% of the total variance, which can be interpreted as an acceptable percentage.

The Cronbach's alpha was calculated for each of the factors obtained in order to estimate the reliability of the results. Factor 1 obtained a Cronbach alpha value of 0.874, Factor 2 obtained a Cronbach alpha value of 0.907 and Factor 3 obtained a Cronbach alpha value of 0.862. This signifies that the values obtained are good, and that the results are therefore reliable.

If we interpret Table III on the basis of the definitions of the DQ characteristics we obtain that the data in Factor 1 must be understandable to users if they are to consider them valid (validity), defined according to regulations and standards (compliance), true and reliable (reliability), contain detailed information to facilitate the task at hand (scope), contain the appropriate and specific information needed for the use to which they will be put (specialization), are verifiable, and the author or source from which they came is known (traceability) and can be adapted to user needs (flexibility). For Factor 2, the users wish to find what they are looking for (effectiveness), with concise data that will allow them to find relevant results (precision), that satisfy their needs (usefulness) and using an appropriate quantity of resources (efficiency). In Factor 3 the users are interested in the fact that the data are applicable, innovative (relevancy) and are oriented towards a destination community (applicability), that they allow them to acquire new knowledge (novelty), to obtain advantages from them (value-added) and that they are available in the shortest possible amount of time (timeliness).

TABLE III.   FACTORIAL ANALYSIS

Factor 1	Factor 2	Factor 3
Validity	Effectiveness	Relevancy
Compliance	Efficiency	Timeliness
Reliability	Precision	Novelty
Scope	Usefulness	Value-added
Specialization		Applicability
Traceability		
Flexibility		

C. Creation of user profiles: Cluster analysis

The DQ characteristics have now been organized into factors. However, our intention was to discover whether any of the DQ characteristics are more relevant than others according to the different user profiles. To do this, we carried out a cluster analysis in order to group the previously identified factors by resemblance or similitude. Three groups are also obtained in this case. Each cluster may be formed of one or various factors, depending on the importance that each cluster gives to each factor. Cluster 1 contains the DQ characteristics of Factor 2. Cluster 2 contains the DQ characteristics of Factor 1 and Factor 3 and Cluster 3 contains the DQ characteristics of Factor 3 (see Table IV).

Our next objective was to determine the relationship between the demographic aspects and each of the clusters identified. This was done by using the contingency tables shown in Table V and the following steps:

1º.- We determined which variable had the greatest value for each demographic aspect and each cluster. For example, Cluster 1 and the demographic aspect 'gender' give us a value of '68' which corresponds with the variable 'male', while the demographic aspect 'level of studies' gives us a value of '54' which corresponds with the variable 'University'.

2º.- The value obtained was compared with the other values in this variable for the other clusters. In the example, the values are '55' and '44' for the variable 'male' in Clusters 2 and 3, respectively, and the values are '50' and '56' for the variable 'University' in Clusters 2 and 3, respectively .

3º.- We chose the greatest of the values obtained for this variable. In the example, we selected the value '68', which is in Cluster 1, for the variable 'male', and the value '56',

which is in Cluster 3, for the variable 'University', and this variable was discounted in Cluster 1 (the values shown in bold type in Table V).

4º.- For those variables which did not yet have a selected value, we chose the highest value in its row. For example, for the 'Vocational Training' variable, whose values are '7', '13', and '20' in Clusters 1, 2 and 3, respectively, we chose the value '20' which is in Cluster 3 (the values shown in italics in Table V). By following these steps we therefore obtain the values shown in bold type and in italics (highlighted values) in Table V.

These results allow each variable of each demographic aspect to be situated in one of the three clusters, which allows three user profiles to be determined, as is shown in Table VI.

The user profile 1 is composed of: men between 25 and 45 with a postgraduate level of studies who work in education belong to Cluster 1 and give priority to the DQ characteristics in Factor 2 (Effectiveness, Efficiency, Precision and Usefulness). In the user profile 2, there are users under 25 and over 55 years of age with High School who belong to an Industrial, Commercial or Financial organization, of which Cluster 2 is composed, give priority to the DQ characteristics that correspond with Factor 1 (Validity, Compliance, Reliability, Scope, Specialization, Traceability and Flexibility) and Factor 3 (Relevancy, Timeliness, Novelty, Value-added and Applicability). The user profile 3 consists of: women between the ages of 45 and 55 with vocational training or university studies from the service sector or another (i.e., not Education, Industrial, Commercial or Financial), belong to the cluster 3 and these users place more relevance on the DQ characteristics in Factor 3 (Relevancy, Timeliness, Novelty, Value-added and Applicability).

TABLE IV.   CLUSTER ANALYSIS

Cluster		
1	2	3
Factor 2	Factor 1	Factor 3
	Factor 3	

TABLE V.   RELATIONSHIP BETWEEN DEMOGRAPHIC ASPECTS AND CLUSTERS

Demographic Aspect	Variable	Cluster (%)		
		1	2	3
Gender	Male	**68**	55	44
	Female	32	45	**56**
Age	Under 25	13	**36**	16
	Between 25 and 35	*32*	23	28
	Between 35 and 45	**39**	7	20
	Between 45 and 55	13	23	**30**
	Over 55	3	*11*	6
Level of Studies	High School	7	**32**	16
	Vocational Training	7	13	*20*
	University	54	50	**56**
	Postgraduate	*32*	5	8
Type of Organization	Education	**58**	41	35
	Industrial-commercial-financial	7	*16*	4
	Service Sector	26	22	*34*
	Other	9	21	*27*

TABLE VI. SUMMARY OF THE RELATIONSHIP BETWEEN DEMOGRAPHIC ASPECTS AND CLUSTER

Profile/ Cluster	Factor	DQ Characteristics	Gender	Age	Level of Studies	Type of Organization
1	Factor 2	Effectiveness, Efficiency, Precision, Usefulness	Male	Between 25 and 35 Between 35 and 45	Postgraduate	Education
2	Factor 1 Factor 3	Validity, Compliance, Reliability, Scope, Specialization, Traceability, Flexibility Relevancy, Timeliness, Novelty, Value-added, Applicability		< 25 > 55	High School	Industrial-commercial-financial
3	Factor 3	Relevancy, Timeliness, Novelty, Value-added, Applicability	Female	Between 45 and 55	Vocational Training University	Services Other

Limitations of the study. This work has been carried out in a systematic manner. Nevertheless, we are conscious that it has certain limitations. The first concerns the quality model used since we have limited the research to DQ characteristics in the Contextual category of the SPDQM model.

We have also limited the type of Web portal with which we have worked, and have focused solely upon 'Information Web portals' and have obtained those DQ characteristics which are most relevant according to the users' different demographic aspects.

In future works we shall consider the other types of Web portals and we shall analyse the other categories in the model.

As will be noted, all these changes will make the results that will be obtained more global, which is our eventual objective.

V. GUIDELINES TO DESIGNERS AND DEVELOPERS

In this section, we show the method used in order to create guidelines for designers and developers so that they will know which DQ characteristics are most important according to the type of user to which the 'Information Web portals' are oriented and which they intend to develop or modify. The designers and developers will therefore be able to deal with and use them.

We used the following steps:

1.- The type of user towards which the Web portal is oriented is identified. The type of user will be determined by the demographic aspects: gender, age range, level of studies and type of organization with which they are linked (for study or work purposes). The following examples will allow us to analyse the results:

Example 1: A 'Information Web portal' oriented principally towards men between 35 and 45 with postgraduate studies and belonging to educational organisations.

Example 2: A 'Information Web portal' oriented principally towards women under 25 with 'High School' studies and belonging to organisations in the services sector.

2.- The cluster and factor belonging to each demographic aspect is obtained (see Table IV and Table VI).

For example 1:
- Men belong to Cluster 1 with Factor 2.
- People between 35 and 45 are in Cluster 1 with Factor 2.
- People with university studies are in Cluster 1 with Factor 2.
- People in educational organisations are in Cluster 1 with Factor 2.

For example 2:
- Women belong to Cluster 3 with Factor 3.

- People under 25 are in Cluster 2 with Factors 1 and 3.
- People with 'High School' studies are in Cluster 2 with Factors 1 and 3.
- People from organisations in the services sector belong to Cluster 3 with Factor 3.

3.- The designers and developers will put special emphasis on the DQ characteristics in the factor that appears most often. If none of the factors are repeated or various factors are repeated the same amount of times, they will consider the DQ characteristics of those factors.

From example 1: emphasis should be placed on the DQ characteristics in Factor 2. Therefore, in 'Information Web portals', their users are interested in data that are actually seeking (Effectiveness), are obtained using appropriate quantity and resources (Efficiency), are exact (Precision) and that satisfy their needs (Usefulness).

From example 2: most importance will be placed on the DQ characteristics from Factor 3. In this case, users are interested in the fact that the data are applicable and innovative (Relevancy), are obtained in the least possible amount of time (Timeliness), that they allow them to acquire new knowledge (Novelty), permit advantages to be attained (Value-added) and data that are oriented towards a destination community (Applicability) in 'Information Web portals'.

We believe that with this type of guidelines, designers and developers will increase the data quality of the Web portals as perceived by their users, since they will be 'tailor-made'.

VI. CONCLUSIONS AND FUTURE WORK

This paper presents a study whose intention is, on the one hand to establish the importance that Web portal users place on a group of DQ characteristics and, on the other to determine whether these DQ characteristics have a different level of relevance according to the various user profiles.

We determined the set of DQ characteristics to be analysed by focusing on those corresponding to the Contextual category in the DQ model for Web portals denominated as SPDQM [20]. Given that the context in which DQ characteristics are analysed has an influence on this Contextual category, we decided to focus our investigation on Web portals of the 'Information Web Portals' type.

Furthermore, the following demographic aspects: gender, age range, level of studies and type of organization to which the users were linked, were used to identify the various user profiles.

The entirety of this study was carried out by using a survey as a starting point and then analysing the results

obtained from the questionnaire by using the SPSS statistical analysis tool. The questionnaire in this survey contained a total of 21 questions, 4 of which were general, related to demographic aspects, 16 of which were related to DQ characteristics in the Contextual category and 1 question concerning the definition of the term 'Contextual'.

In the analysis that followed we identified three different user profiles and we grouped the DQ characteristics according to those user profiles such that we are now able to indicate the set of DQ characteristics to which more attention should be paid in the DQ of a Web portal according to a particular user profile.

All of this should enable designers and developers to be guided in the construction of Web portals, in order to make them more appropriate for the users at which they are aimed.

Our short-term future work will be to determine the importance of DQ characteristics in the Contextual category for Web portals of the types 'Commercial Interaction' and 'Interaction with other People'. We shall then compare our current results with the results obtained in the other types of portals. In this way, we will see whether all the types of user profiles place importance on the same DQ characteristics in all types of portals. Our eventual intention is to make our model available to users and developers through a free tool.

REFERENCES

[1] M. Komathi and I. Maimunah, "Influence of gender role on Internet usage pattern at home among academicians," The Journal of International Social Research, vol. 2, 2009.

[2] K. Laudon and C. Traver, E-commerce 2009 (5th ed.): Upper Saddle River, NJ: Prentice Hall, 2008.

[3] P. Sharma and J. Gupta, "A framework for enterprise-wide-e-commerce portal for evolving organizations," in A. Tatnall (Ed.), Web portals: The new gateways to internet information and services. Hershey, PA: Idea Group ed, 2005.

[4] R. Stair and G. Reynolds, Principles of information systems.: Boston, MA: Thomson Course Technology, 2008.

[5] A. S. Al-Mudimigh, Z. Ullah and T. A. Alsubaie, "A framework for portal implementation: A case for Saudi organizations," International Journal of Information Management, vol. 31, pp. 38-43, 2011.

[6] P. Yang, et al., "The emerging concepts and applications of the spatial web portal," Photogrammetric Engineering & Remote Sensing, vol. 73, pp. 691-698, 2007.

[7] M. A. Domingues, C. Soares and A. M. Jorge, "A Web-Based System to Monitor the Quality of Meta-Data in Web Portals," in IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IATW'06), 2006, pp. 188-191.

[8] H. Collins, "Corporate Portal Definition and Features," AMACOM, 2001.

[9] M. Davydov, Corporate Portals and e-Business Integration Emerging Business Technology Series, 2001.

[10] C.-T. Liu, T. C. Du and H-H. Tsai, "A study of the service quality of general portals," Information & Management, vol. 46, pp. 52-56, 2009.

[11] A. Durndell and Z. Haag, "Computer self efficacy, computer anxiety, attitudes towards the Internet and reported experience with the Internet, by gender, in an East European sample," Computer in Human Behavior, vol. 18, pp. 521-535, 2002.

[12] M. E. Hupfer and B. Detlor, "Gender and Web information seeking: A self-concept orientation model.," Journal of the American Society for Information Science and Technology, vol. 57, pp. 1105-1115, 2006.

[13] D. Kim, X. Lehto and A. Morrison, "Gender differences in online travel information search: Implications for marketing Communications on the Internet," Tourism Management Information Systems; Armonk; Spring, vol. 28, pp. 423-433, 2007.

[14] C. Şahin, "An analysis of Internet addiction levels of individuals according to various variables," TOJET: The Turkish Online Journal of Educational Technology, vol. 10, pp. 60-66, 2011.

[15] C. Cappiello, C. Francalanci and B. Pernici, "Data quality assessment from the user´s perspective," in Proceeding on International Workshop on Information Quality in Information Systems (IQIS2004), Paris, France. ACM, 2004, pp. 68-73.

[16] D. Strong, Y. Lee and R. Wang, "Data Quality in Context," Communications of the ACM, vol. 40, pp. 103-110, 1997.

[17] L. Pipino, Y. Lee and R. Wang, "Data Quality Assessment," Communications of the ACM, vol. 45, pp. 211-218, 2002.

[18] R. Wang and D. Strong, "Beyond accuracy: What data quality means to data consumers," Journal of Management Information Systems; Armonk; Spring, vol. 12, pp. 5-33, 1996.

[19] M. Gertz, T. Ozsu, G. Saake and K-U. Sattler, "Report on the Dagstuhl Seminar "Data Quality on the Web"," SIGMOD Record., vol. 33, pp. 127-132, 2004.

[20] C. Moraga, M. Moraga, C. Calero and A. Caro, "SQuaRE-Aligned Data Quality Model for Web Portals," in 9th International Conference on Quality Software (QSIC 2009), 2009, pp. 117-122.

[21] B. Kitchenham and S. Pfleeger, "Principles of survey research part 2: designing a survey," SIGSOFT: Software Engineering Note, vol. 27, pp. 18-20, 2002.

[22] B. Kitchenham and S. Pfleeger, "Principles of survey research part 3: Constructing a survey instrument," SIGSOFT: Software Engineering Note, vol. 27, pp. 20-24, 2002.

[23] M. J. Norusis, SPSS 11.0 Guide to Data Analysis: Prentice Hall, 2002.