

IADIS International Conference

WWW/INTERNET

18 - 21 October @ Madrid - Spain

2012

Proceedings

Edited by:
Bebo White and
Pedro Isaías

IADIS INTERNATIONAL CONFERENCE

WWW/INTERNET 2012

**PROCEEDINGS OF THE
IADIS INTERNATIONAL CONFERENCE
WWW/INTERNET 2012**

MADRID, SPAIN

OCTOBER 18-21, 2012

**Organised by
IADIS**

International Association for Development of the Information Society

Copyright 2012

IADIS Press

All rights reserved

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Permission for use must always be obtained from IADIS Press. Please contact secretariat@iadis.org

Edited by Bebo White and Pedro Isaías

Associate Editor: Luís Rodrigues

ISBN: 978-989-8533-09-8

TABLE OF CONTENTS

FOREWORD	xi
PROGRAM COMMITTEE	xv
KEYNOTE LECTURES	xix
FULL PAPERS	
SEMANTIC RETRIEVAL OF DOCUMENTS FROM DIGITAL REPOSITORIES IN THE MOODLE ENVIRONMENT <i>Renan Rodrigues de Oliveira, Fábio Moreira Costa, Cedric Luiz de Carvalho and Ana Paula Ambrósio</i>	3
HTML SEGMENTATION USING ENTROPY GUIDED TRANSFORMATION LEARNING <i>Evelin Carvalho Freire de Amorim</i>	11
WPPS: A NOVEL AND COMPREHENSIVE FRAMEWORK FOR WEB PAGE UNDERSTANDING AND INFORMATION EXTRACTION <i>Ruslan R. Fayzrakhmanov</i>	19
AN APPROACH FOR EXTRACTING WEB FORM LABELS BASED ON DISTANCE ANALYSIS OF HTML COMPONENTS <i>Leonardo Bres dos Santos, Carina F. Dorneles and Ronaldo dos S. Mello</i>	27
EXTRACTING AND EXPOSING RELATIONAL DATABASE METADATA ON THE WEB <i>João Carlos da Silva, Elisabete Tomomi Kowata and Auri Marcelo Rizzo Vincenzi</i>	35
BUSINESS MODELS FOR MOBILE APPLICATIONS <i>Alexandra Chapko, Andreas Emrich, Marc Gräßle, Dirk Werth and Peter Loos</i>	43
AN INFLUENCE PERSPECTIVE: IS USER PARTICIPATION CRUCIAL IN THE WEB DEVELOPMENT PROCESS? <i>Tomayess Issa and Pedro Isaías</i>	51
AWARENESS OF OTHERS IN ACCESSIBLE COLLABORATIVE RICH INTERNET APPLICATIONS <i>Leonelo D. A. Almeida and M. Cecília C. Baranauskas</i>	60
DATA QUALITY IN WEB PORTALS FOR INTERACTION WITH OTHER PEOPLE <i>Carmen Moraga, M^a Angeles Moraga, Angélica Caro, Rodrigo Romo Muñoz and Coral Calero</i>	68

AUTHORITYRANK: COGNITIVE AUTHORITY AND INFORMATION RETRIEVAL IN THE WEB	76
<i>Filipe Roseiro Cogo, Sérgio Roberto P. da Silva and Roberto Pereira</i>	
A WEB-BASED METHOD FOR ONTOLOGY POPULATION	85
<i>Hilário Oliveira, Rinaldo Lima, João Emanuel and Fred Freitas</i>	
DOMAIN ONTOLOGIES MODELING VIA SEMANTIC ANNOTATIONS OF UNSTRUCTURED WIKI KNOWLEDGE	93
<i>Roberto Boselli, Mirko Cesarini, Fabio Mercorio and Mario Mezzanatica</i>	
SOCIAL LEARNING: DEFINING LEARNING OBJECTS FROM SOCIAL TOOL	101
<i>André Luís Andrade Menolli, Sheila Reinehr and Andreia Malucelli</i>	
A UBIQUITOUS ELECTRONIC TOURIST GUIDE FOR THE CAMINHOS DE PEDRA ITINERARY	109
<i>Humberto Costa, Cristiano Costa, Eduardo Silva, Sandro Rigo, Jorge Barbosa, Luiz Silveira Jr. and Underlea Bruscato</i>	
RECO – A FRAMEWORK FOR EXPERIMENTATION WITH RECOMMENDERS	117
<i>Jakub Ševcech, Michal Kompan and Mária Bielíková</i>	
TRANSLATING XML QUERIES INTO EQUIVALENT SQL STATEMENTS	125
<i>Marta Breunig Loose and Deise de Brum Saccol</i>	
A SURVEY ON SOCIAL NETWORK SITES' FUNCTIONAL FEATURES	133
<i>Antonio Tapiador and Diego Carrera</i>	
A STUDY OF THE CONDITIONS FOR A GOOD DIGITAL CITIZEN IN A NEW MEDIA ERA	141
<i>Heein Yang, Misoo Kwon, Gilwoo Nam and Jongsoo Jeon</i>	
WEB-BASED INFORMATION EXPLORATION OF SENSOR WEB USING THE HTML5/X3D INTEGRATION MODEL	149
<i>Byounghyun Yoo</i>	
A RUBY DOMAIN SPECIFIC LANGUAGE (DSL) FOR WEB MASHUPS	157
<i>Richard J Barnett and Carianne Cowley</i>	
A REAL-TIME WEB-BASED HEALTH MONITORING SYSTEM BASED ON ENTERPRISE SERVICE BUS	165
<i>Yohanes Baptista Dafferianto Trimugroho, Kamyar Rasta, Trinh Hoang Nguyen, Rune Fensli and Frank Reichert</i>	
HOW CAN A MOBILE VENDOR ENGENDER SHOPPER TRUST AND REDUCE PERCEIVED OPPORTUNISM?	173
<i>Sonia San-Martín</i>	
INFRASTRUCTURE TO NEXT-GEN PROACTIVE E-COMMERCE ENVIRONMENT THROUGH INTERNET: UBIQUITOUS COMMERCE FOR THE MASSES	180
<i>Paulo H. Cazarotto, Cristiano André da Costa, Rodrigo da Rosa Righi and Jorge L. V. Barbosa</i>	
ABANDONMENT IN WEB APPLICATIONS FOR PURCHASING AIRLINE TICKETS	188
<i>Malcolm Mactavish and Lyndon Veale</i>	
CYBERSQUATTING DETECTION BY AUTOMATIC DOMAIN NAME RE-ACCENTING	195
<i>Jean Ceccato and Anthony Don</i>	

WHAT DETERMINES E-LOYALTY? AN ANALYSIS OF FACTORS AFFECTING ON-LINE CUSTOMER RETENTION <i>Donnacha Clifford and Michael Lang</i>	203
INFLUENCE OF PERCEIVED QUALITY OF UNIVERSITY OFFICIAL WEBSITE TO PERCEIVED QUALITY OF UNIVERSITY EDUCATION AND ENROLLMENT INTENTION <i>Achmad Nizar Hidayanto, Fanny Rofalina and Putu Wuri Handayani</i>	211
AUTOMATIC AND CONTINUOUS MONITORING AND COMPOSITION OF NONDETERMINISTIC WEB SERVICES <i>Gustavo B. Valfre, Angelo E. M. Ciarlini and Sean W. M. Siqueira</i>	219
USABILITY EVALUATION OF ELECTRONIC SIGNATURE BASED E-GOVERNMENT SOLUTIONS <i>Thomas Zefferer and Vesna Krnjic</i>	227
A SEMANTIC WEB APPROACH FOR AUTOMATED TEST GENERATION <i>Bogdan Drăgulescu, Marian Bucos and Radu Vasiu</i>	235
EXPLORING PARTICIPATORY DESIGN FOR SNS-BASED AEH SYSTEMS <i>Lei Shi, Dana Al Qudah and Alexandra I. Cristea</i>	242
DEAF LITERACY: A COMPUTATIONAL PROCESS TO DESIGN SIGN LANGUAGE/PORTUGUESE ARTIFACTS FOR INTERNET <i>Cayley Guimarães, Diego R. Antunes, Laura S. Garcia, Leticia M. Peres and Sueli Fernandes</i>	250
PARALLEL HIGHER-ORDER SVD FOR TAG-RECOMMENDATIONS <i>Philipp Shah, Christoph Wieser and François Bry</i>	257
A TRAVEL SEQUENCE RECOMMENDATION APPROACH BASED ON MARKOV MODEL <i>Dalu Guo, Richong Zhang, Xudong Liu and Hailong Sun</i>	266
LEARNING SYNONYM RELATIONS FROM FOLKSONOMIES <i>Alex Rêgo, Leandro Marinho and Carlos Eduardo Pires</i>	273
EXPLORING THE EFFECTS OF SOCIAL INFLUENCE ON USER BEHAVIOR TARGETED TO FEEDBACK SHARING <i>Agnis Stibe and Harri Oinas-Kukkonen</i>	281
USABILITY EVALUATION OF FACEBOOK'S PRIVACY FEATURES: COMPARISON OF EXPERTS AND USERS <i>Marcos H. Kimura, Marcio J. Mantau, Avaniilde Kemczinski, Isabela Gasparini and Carla D. Medeiros Berkenbrock</i>	290
SUBJECT CLASSIFICATION OF WEB PAGES <i>Ludger Martin</i>	298
USATASKER: A TASK DEFINITION TOOL FOR SUPPORTING THE USABILITY EVALUATION OF WEB APPLICATIONS <i>Leandro Guarino de Vasconcelos and Laércio Augusto Baldochi Jr.</i>	307
ASSESSING THE PERFORMANCE OF JAVA AND ERLANG IN WEB 2.0 APPLICATIONS <i>Jucimar Maia da Silva Jr., Rafael Dueire Lins and Lanier Menezes dos Santos</i>	315

EMULATION OF COMPLEX NETWORK INFRASTRUCTURES FOR LARGE-SCALE TESTING OF DISTRIBUTED SYSTEMS <i>Robert Lübke, Robin Lungwitz, Daniel Schuster and Alexander Schill</i>	323
A CLIQUE BASED WEB GRAPH MODEL <i>Zhiming Chen and Andrés Faragó</i>	331
INTEGRATING PROCESS AND SERVICES THROUGH META-MODELS <i>Patricia Bazán, Roxana Giandini, Gabriela Perez, Elsa Estevez and Javier Diaz</i>	339
AN ADAPTIVE APPROACH FOR IDENTIFYING REPUTATION OF RESEARCHERS <i>Cristiano Roberto Cervi, Renata Galante and José Palazzo Moreira de Oliveira</i>	347
STRATEGIES AND MOTIVATIONS BEHIND ARTIFICIAL TRENDING TOPICS IN TWITTER <i>Raquel Recuero and Ricardo Araujo</i>	355
GUIDE OF ACCESSIBILITY AND USABILITY RECOMMENDATIONS AIMING FOR THE DEVELOPMENT OF HYPERMEDIA FOR DEAF <i>Vanessa Tavares de Oliveira Barros, Kátia Tavares Meserlian, Rodolfo Miranda de Barros, Flávio Ogawa and Francisco Antônio Pereira Fialho</i>	363
COMPARISON OF USABILITY TESTING TOOLS FOR WEB GRAPHICAL INTERFACES <i>Bernardo Santos, Carlos Teixeira and Ana Respício</i>	371
IDENTIFYING PRAGMATIC PATTERNS OF COLLABORATIVE PROBLEM SOLVING <i>Heiko Hormung, Roberto Pereira, M. Cecilia C. Baranauskas, Rodrigo Bonacin and Julio Cesar Dos Reis</i>	379
INTERPRETATION OF WEB SITE USER INTERACTION AS A BASE FOR CONTEXT-AWARE PAGE ADAPTATION <i>Pascal Bihler</i>	388

SHORT PAPERS

HEURISTIC AND AI APPROACH TO OPTIMIZE PLAGIARISM DETECTION TOOL USING A PUBLIC SEARCH ENGINE <i>Ondrej Vesely, Jan Kolomaznik and Tomas Foltyněk</i>	399
ONTOLOGY ON THE LEVEL G OF THE SOFTWARE PROCESS MODEL MPS.BR TO ASSIST BUSINESS PROCESSES MODELING <i>Alessandro Viola Pizzoleto, Hilda Carvalho de Oliveira and Celso Socorro Oliveira</i>	404
A MULTI-TOOL SCHEME FOR SUMMARIZING TEXTUAL DOCUMENTS <i>Rafael Dueire Lins, Steven J. Simske, Luciano de Souza Cabral, Gabriel de França Silva, Rinaldo Lima, Rafael F. Mello and Luciano Favaro</i>	409
AUTOMATION OF A VENDING SYSTEM USING SMARTPHONES <i>Tiago Jost, Cristiano Costa, Rodrigo Righi and Alexandre Andrade</i>	415

RERANKING IMAGE SEARCH RESULT BASED ON PHOTOGRAPHIC QUALITY ASSESSMENT WITHOUT IMAGE FEATURES <i>Masaharu Hirota, Shohei Yokoyama, Naoki Fukuta and Hiroshi Ishikawa</i>	420
TOPIC PAGE MINING BASED ON PHRASERANK FOR ADVERTISEMENT IMAGE <i>Jian Sun, Siyuan Chen, Yingju Xia and Jun Sun</i>	425
ON THE USE OF SOCIAL MEDIA FOR IMPORTANT THINGS: FACEBOOK AS A COPING TOOL <i>Felix S. Hussenoeder</i>	431
GENAPI: A GENERIC SOCIAL-NETWORKING API <i>Richard J. Barnett and Simon de la Rouviere</i>	436
DISTINCTION BETWEEN OPINION AND INFORMATION SPREADING IN SOCIAL NETWORKS <i>Igor Kanovsky and Omer Yaari</i>	442
A MODELING APPROACH FOR KNOWLEDGE MANAGEMENT IN COMPLEX BUSINESS SYSTEMS <i>Abdussalam Ali and Igor Hawryszkiewwycz</i>	447
AN INCREMENTAL APPROACH TO TECHNOLOGY-SUPPORTED CURRICULUM DESIGN AND APPROVAL <i>Jim Everett, George Macgregor and Rehman Mohamed</i>	453
MEANINGFUL LEARNING IN MATHEMATICS EDUCATION: A PROPOSAL OF DEVELOPING A PROTOTYPE OF AN AUGMENTED REALITY TOOL TO SUPPORT THE TEACHING OF CALCULATION <i>Edson Pacheco and Rodolfo Miranda de Barros</i>	458
SOCIAL NETWORK ENGINEERING AND ONTOLOGY ENGINEERING FOR E-LEARNING: HOW DO THESE WORK TOGETHER? <i>Seid Maglajlic</i>	463
DETERMINATION OF TOPIC DESCRIPTION TERMS IN TOPIC MODELS <i>Kota Kamura, Hung-Hsuan Huang and Kyoji Kawagoe</i>	468
METROPOLITALIA: A CROWDSOURCING PLATFORM FOR LINGUISTIC FIELD RESEARCH <i>Fabian Kneissl and François Bry</i>	473
INVESTIGATING COLLABORATION AND EFFECTIVENESS OF VIRTUAL TEAMS WITH DISTINCT ORGANIZATION TYPES <i>Juliana de Melo Bezerra, Celso Massaki Hirata and André Antônio Battagello</i>	477
WEB ANALYTICS AS ONE OF THE FEEDBACK MECHANISMS IN ELECTRONIC GOVERNMENT MANAGEMENT <i>Rasim Alguliev, Hana Skalska, Farhad Yusifov and Mahammad Sharifov</i>	483
BRICK: A LINKED DATA EXPERIENCE <i>Nadia Catenazzi, Lorenzo Sommaruga and Ingrid Domenighetti</i>	487
LESSONS LEARNED FROM CREATING A TRUST SYSTEM FOR P2P MARKETPLACES <i>Mauro Nunes and João Correia</i>	492

REFLECTION PAPERS

EDUCA REPOSITORY SERVICE: API TO SUPPORT DIFERENTES DIGITAL REPOSITORIES <i>Jorge Rocha, Ana Caixinha, Joaquim Arnaldo Martins and Marco Fernandes</i>	499
INTERACTION DESIGN ISSUES FOR MOBILE MULTI-TOUCH APPS <i>Huberta Kritzenberger</i>	503
BUILDING TEST SUITES FROM TEST RECORDINGS OF WEB APPLICATIONS: REFLECTION PAPER <i>Martin Filipisky, Miroslav Bures and Ivan Jelinek</i>	507
PEOPLE'S HISTORY APPLICATIONS: USING THE WWW TO RE-WRITE HISTORY <i>Lorie Loeb and Sang Jin Lee</i>	511
DATA PROTECTION AND EMPLOYEE BEHAVIOUR: THE ROLE OF INFORMATION SYSTEMS SECURITY CULTURE <i>Lena Connolly and Michael Lang</i>	516

POSTERS

E-LEARNING IN INFORMATICS TEACHING <i>Olga Mironova, Irina Amitan and Jüri Vilipõld</i>	523
SOCIAL CHEESECAKE: AN UX-DRIVEN INTERFACE FOR MANAGING CONTACTS <i>Alicia Diez and Antonio Tapiador</i>	526
DESCRIPTION STANDARDS: CROSSWALK PROPOSAL FOR EDUCA <i>Ana Caixinha, Joaquim Arnaldo Martins, Jorge Rocha and Marco Fernandes</i>	529
CAREER EDUCATION SUPPORT UTILIZING MOBILE DEVICES <i>Rumiko Kurita, Motoo Kusano, Asahiro Arai, Shoichi Nakamura, Setsuo Yokoyama and Youzou Miyadera</i>	533
TOWARDS AN APPROACH BASED ON ELECTRONIC CONTRACT TO ADDRESS THE VENDOR LOCK-IN IN CLOUD COMPUTING <i>Gabriel Costa Silva and Itana Maria de Souza Gimenes</i>	537

AUTHOR INDEX

FOREWORD

These proceedings contain the papers and posters of the IADIS International Conference WWW/Internet 2012, which was organised by the International Association for Development of the Information Society, Madrid, Spain, 18-21 October 2012.

The IADIS WWW/Internet 2012 Conference aims to address the main issues of concern within WWW/Internet. WWW and Internet had a huge development in recent years. Aspects of concern are no longer just technical anymore but other aspects have arisen. This conference aims to cover both technological as well as non-technological issues related to these developments.

Submissions were accepted under the following main tracks and topics:

- Web 2.0
 - Collaborative Systems
 - Social Networks
 - Folksonomies
 - Enterprise Wikis and Blogging
 - Mashups and Web Programming
 - Tagging and User Rating Systems
 - Citizen Journalism

- Semantic Web and XML
 - Semantics Web Architectures
 - Semantic Web Middleware
 - Semantic Web Services
 - Semantic Web Agents
 - Ontologies
 - Applications of Semantic Web
 - Semantic Web Data Management
 - Information Retrieval in Semantic Web

- Applications and Uses
 - e-Learning
 - e-Commerce / e-Business
 - e-Government
 - e-Health
 - e-Procurement
 - e-Society
 - Digital Libraries
 - Web Services/SaaS
 - Application Interoperability
 - Web-based multimedia technologies

➤ Services, Architectures and Web Development

- Wireless Web
- Mobile Web
- Cloud/Grid Computing
- Web Metrics
- Web Standards
- Internet Architectures
- Network Algorithms
- Network Architectures
- Network Computing
- Network Management
- Network Performance
- Content Delivery Technologies
- Protocols and Standards
- Traffic Models

➤ Research Issues

- Web Science
- Digital Rights Management
- Bioinformatics
- Human Computer Interaction and Usability
- Web Security and Privacy
- Online Trust and Reputation Systems
- Data Mining
- Information Retrieval
- Search Engine Optimization

The IADIS WWW/Internet 2012 Conference had 201 submissions from more than 33 countries. Each submission has been anonymously reviewed by an average of four independent reviewers, to ensure the final high standard of the accepted submissions. The final result was the approval of 49 full papers, which means that the acceptance rate was 25%. A few more papers have been accepted as short papers, reflection papers, and poster/demonstrations. Best papers will be selected for publishing as extended versions in the IADIS International Journal on WWW/Internet (IJWI) and in other selected journals.

The conference, besides the presentation of full papers, short papers, reflection papers, and posters presentations also included two keynote presentations from internationally distinguished researchers. We therefore would like also to express our gratitude to Dr. Irwin King, The Chinese University of Hong Kong, Hong Kong and Dr. Ricardo Baeza-Yates, VP of Yahoo! Research, Barcelona, Spain.

As we all know, organising a conference requires the effort of many individuals. We would like to thank all members of the Program Committee for their hard work in reviewing and selecting the papers that appear in the proceedings.

We are especially grateful to the authors who submitted their papers to this conference and to the presenters who provided the substance of the meeting.

These Proceedings book contain a rich experience of the academic & research institutions and the industry on diverse themes related to the Internet and Web. We do hope that researchers, knowledge workers and innovators both in academia and the industry will find it a valuable reference material.

Last but not the least, we hope that everybody will have a good time in Madrid and we invite all participants for the next year's edition of the IADIS International Conference WWW/Internet that will be held in Fort Worth, Texas, USA.

Bebo White, Stanford University, USA
Program Chair

Pedro Isaiás, Universidade Aberta (Portuguese Open University), Portugal
Conference Chair

Madrid, Spain
18 October 2012

PROGRAM COMMITTEE

PROGRAM CHAIR

Bebo White, Stanford University, USA

CONFERENCE CHAIR

Pedro Isafas, Universidade Aberta (Portuguese Open University), Portugal

COMMITTEE MEMBERS

Abdolhossein Sarrafzadeh, UNITEC New Zealand, New Zealand
Agostino Poggi, Universitat Degli Studi Di Parma, Italy
Alberto Barrón-Cedeño, Universitat Politècnica de Catalunya, Spain
Alessandro Soro, Crs4, Italy
Alexander Schill, TU Dresden, Germany
Alexander Trousov, IBM Dublin Center for Advanced Studies, Ireland
Alexander Woehrer, St. Poelten University of Applied Sciences, Austria
Alexiei Dingli, University Of Malta, Malta
Ana Fermoso, Universidad Pontificia De Salamanca, Spain
Ana Sanz, Universidad Carlos III De Madrid, Spain
Ana Regina Cavalcanti Da Rocha, Centroin, Brazil
Ananya S. Guha, Indira Gandhi National Open University, India
Andrea Kienle, University of Applied Sciences, Dortmund, Germany
Andreas Papasalouros, University Of The Aegean, Greece
Andreas Schrader, Insitute of Telematics, University of Luebeck, Germany
Angélica Antonio, Universidad Politécnica De Madrid, Spain
Anirban Kundu, West Bengal University Of Technology, India
Antonio Reyes, Instituto Superior de Intérpretes y Traductores, Mexico
Arturo Mora-Soto, Carlos III University Of Madrid, Spain
Asadullah Shaikh, Universitat Oberta de Catalunya, Spain
Beob Kyun Kim, Korea Institute of Science and Technology Informat, South Korea
Bo Hu, Universität Der Bundeswehr München, Germany
Brahmananda Sapkota, University Of Twente, Netherlands
Carsten Ulrich, Shanghai Jiao Tong University, China
Charlie Abela, University Of Malta, Malta
Chih Cheng Hung, Southern Polytechnic State University, USA
Christos Bouras, University Of Patras, Greece
Christos Georgiadis, University Of Macedonia, Thessaloniki, Greece
Christos Katsanos, Hellenic Open University, Greece
Ciprian Dobre, University Politehnica Of Bucharest, Romania
Clodoveu Davis, Universidade Federal de Minas Gerais, Brazil
Constantina Costopoulou, Agricultural University Of Athens, Greece
Costas Yialouris, Agricultural University Of Athens, Greece
Daniel Cunliffe, University Of Glamorgan, United Kingdom

DATA QUALITY IN WEB PORTALS FOR INTERACTION WITH OTHER PEOPLE

Carmen Moraga¹, M^a Ángeles Moraga¹, Angélica Caro², Rodrigo Romo Muñoz² and Coral Calero¹

¹University of Castilla-La Mancha, Paseo de la Universidad 4, Ciudad Real, Spain

²University of Bio Bio, Chillán, Chile

ABSTRACT

Web portals for interaction with other people are being used on an ever-increasing basis. However, there is a great deal of competition to stay in the market and obtain user-loyalty. In this respect, we consider that the quality of the data provided by these Web portals may influence the opinion that their users have of them. Moreover, aspects such as gender and age range may influence the aspects that are believed to be most important when evaluating the quality of the portal. In this paper, we attempt to analyse this perception. We have therefore selected a set of data quality characteristics with which we shall work and which are related to the quality of this type of portals. We have then created surveys with which to discover users' opinions of these characteristics, and we have analysed the results obtained. Finally, we have studied the influence of the demographic aspects of 'gender', 'age range', 'level of studies' and 'type of organisation to which the users are linked' to see which data quality characteristics are most important according to these aspects. An example of the results obtained is that, men place more importance on data which are oriented towards a destination community (Applicability), whilst women are more concerned that data should be reliable (Reliability).

KEYWORDS

Data quality, web portal, user preferences.

1. INTRODUCTION

A Web portal is a gateway to information and services from various sources that integrates content, and other software services or applications (Tatnall, 2005, Averweg, 2007). After their appearance and commercial growth in 2000 (Manouselis et al., 2009), Web portals have lately attracted increased research interest that focuses on a variety of aspects such as their business models, interface design, technical development, or their quality (Mahadevan, 2000, Tatnall, 2007).

Depending on the principal kind of activity carried out in them, Web portals can be classified in the following types:

- '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': these portals are 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.

In this paper, we shall focus on the last type of Web portals, with the objective of analysing the importance of Data Quality (DQ) characteristics from the point of view of users of this type of Web portals.

We shall also study whether the users' gender, age range, level of studies or the type of organisation to which they belong influences the importance that they place on the DQ characteristics studied, and whether it is possible to create different user profiles.

The remainder of this paper is organized as follows: Section 2 describes the characteristics of the DQ characteristics that will be studied. Section 3 describes the method used to determine the importance of the

DQ characteristics and the creation of user profiles according to the demographic aspects and user preferences. Section 4 shows an analysis of the users according to the various demographic aspects. In Section 5, the preferences of the different user profiles are obtained as regards the DQ characteristics for Web portals of the 'Interaction with other people' type. Finally, Section 6 presents our conclusion and future works.

2. CHARACTERISTICS TO BE ANALYSED

Owing to the importance of the DQ in Web portals, in (Moraga et al., 2009), we defined a DQ model, denominated as SPDQM (SQuaRE-Aligned Portal Data Quality Model). This DQ model is made up of 42 DQ characteristics which are classified in four categories (Intrinsic, Contextual, Representational and Operational). In this paper we focus on those in the Contextual category of the model. This category contains 10 DQ characteristics and 6 DQ sub-characteristics (see Figure 1).



Figure 1. DQ characteristics in the contextual category

3. METHOD USED TO STUDY THE DQ AND USERS OF WEB PORTALS

Our objective is to discover what opinion the users of Web portals of the 'Interaction with other people' type have as regards the DQ characteristics mentioned in Section 2, according to the following demographic aspects: gender, age range, level of studies and type of organisation to which the users are linked.

The method used to discover users' opinions was that of compiling information through the use of surveys. The surveys were carried out on the basis of the work of (Kitchenham and Pfleeger, 2002). They were of the non-supervised type and were distributed to a heterogeneous group of users. In our case, we distributed the surveys by e-mail or in printed format to a total of 200 people from Europe and Latin America. The surveys 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%.

3.1 Questions in the Survey

The questions in the survey were divided into two types: those related to users (see Table 1), and those related to the users' opinions of the DQ characteristics (see Table 2).

The questions concerning the DQ characteristics had to be easy to understand. Pre-test surveys were 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. The survey included 17 questions, 16 of which were related to the DQ characteristics and one of which was created to discover the users' opinions of the definition of the 'Contextual' category itself (the category to which the DQ characteristics belong). Each question was responded with a value between (0) – 'totally disagree' and (10) – 'totally agree'. This scale was selected in order to better refine the results and improve the analysis of the data obtained from each question.

Table 1. Questions concerning demographic aspects.

Gender: Male/Female
Level of studies COMPLETED: Primary Education Qualification / Secondary Education Qualification / Vocational Training/ University/ Post Graduate.
Type of organization to 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 / Industrial / Commercial / Service Sector / Financial / Other (Please state which).
Age range: Under 25 / Between 25 and 35 / Between 35 and 45 / Between 45 and 55 / Between 55 and 65 / over 65.

Table 2. Some questions in the questionnaire for 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).
6.- The data provided by a Web portal should contain the appropriate and specific information for the use to which they will be put.
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.

3.2 Sample Analysis

Having collected the surveys we then analysed the data. The sample is first analyzed according to the questions related to the demographic aspects (Tabla 3).

Table 3. Sample

Demographic Aspect	Variable	Percentage (%)	Total (%)
Gender	Male	50.5	100
	Female	49.5	
Age Range	Under 25	20.2	100
	Between 25 and 35	27.7	
	Between 35 and 45	20.2	
	Between 45 and 55	25.5	
	Between 55 and 65	5.9	
	Over 65	0.5	
Level of Studies	Primary/Secondary Education	18	100
	Vocational Training	16	
	University	54.3	
	Postgraduate	11.7	
Type of Organization to which the users are linked	Education	40.4	100
	Industrial	1.6	
	Commercial	2.1	
	Service Sector	30.3	
	Financial	3.7	
	Other	21.9	

This table shows that the percentage of users is practically the same as regards gender, since under average equality conditions both genders access the Internet indistinctly. Moreover, as (Komathi and Maimunah, 2009) state “with the development of the Internet, both male and female users to certain extent are unquestionably dependent on Internet as a medium of sharing ideas, building communication network as well as for information search “.

As regards age, in our case all the ranges of users under 55 are representative. However, users of 55 or over are far less representative. This is coherent as regards use of the Internet since its use usually decreases as age increases, and user’s preferences are different depending on the age (Makesrithongkum, 2009). But larger percentages of older generations are online now than in the past (Jones and Fox, 2009). In order to make the group more representative we decided to place the last two age ranges in one single group (users over 55) with a total representation of 6.4%.

As regards the level of studies and the type of organisation, there is a greater representation of users with university studies and who are in organisations linked to ‘Education’. This is positive because Web portals of the ‘Interaction with other people’ type are widely used in the sphere of education since, as (Komathi and Maimunah, 2009) state “Internet has become an increasingly important tool in academic careers”. What is more, (Şahin, 2011) points out that the level of studies influences the use of Web portals since “there is a significant difference between the internet addiction scores of students and other professional groups”.

Finally, we should point out that in the 'Type of organisation to which the users are Linked' demographic aspect, the representation of the 'Industrial', 'Commercial' and 'Financial' groups is very low, and from here on they will therefore be grouped as a single variable.

4. ANALYSIS OF RESULTS

Having analysed the sample, we shall now study the results as regards the DQ characteristics. The reliability of the results will first be analysed by using the Cronbach's alpha, with which a highest value of 0.6 is considered to be adequate. In our case, this value is 0.953, and the study therefore has good internal consistency and the results are reliable. Having confirmed the reliability of the sample we can therefore proceed to actually analyse the data, a task that we decided to tackle from three different perspectives by carrying out three different types of analyses.

- **Calculation of descriptive statistics:** This is oriented towards obtaining the minimum and maximum values and the central tendency (mean) of each of the DQ characteristics in our study. The objective of carrying out this analysis is to verify that the characteristics that we have identified are important from the user's point of view.

- **Factorial analysis:** The initial set of DQ characteristics is used to create homogeneous groups of characteristics (denominated as factors). These factors summarise and synthesise the information since they reduce the quantity of initial DQ characteristics. These groups are formed of those DQ characteristics that have a considerable correlation with others. What is more, each group should be independent of the others. In this case the intention is to search for the existence of groups of DQ characteristics, which could later be characterised through the use of the demographic data.

- **Cluster analysis:** The scores obtained from the factorial analysis are used to construct groups of users (denominated as clusters). These clusters determine the importance placed on the DQ characteristics of each factor. The clusters are next related to the demographic aspects, and user profiles are then created according to the results obtained. Finally, the DQ characteristics which are most important to the different user profiles obtained are placed in groups.

4.1 Descriptive Statistical Analysis

A descriptive analysis is first carried out to obtain the minimum, maximum and mean values of each of the DQ characteristics that we wish to study. The mean value of all these DQ characteristics is between 6.57 and 7.35. It can therefore be deduced that all of them are of approximately the same importance for users of Web portals of the 'Interaction with other people' type. Within this margin, the lowest value is 6.57 for 'Novelty' (new data) and the highest value is 7.35 for 'Effectiveness' (data which are those that the user is actually seeking). This indicates that in this type of Web portals the users are less concerned that the data are new to them, and are more preoccupied with obtaining the data that they are actually seeking.

4.2 Factorial Analysis

We shall now carry out a factorial analysis, since it is a data reduction technique which allows the maximum amount of information to be explained with the minimum number of dimensions. In our case, three homogeneous groups of DQ characteristics (denominated as factors) have been determined, signifying that there is a correlation with a minimum value of 0.5 between the DQ characteristics in each group, and each group is in turn independent (see Table 4).

