2017“互联网+、大数据与商业创新国际会议”论文集

冷柏军　罗　欣　主编

对外经济贸易大学出版社

2017“互联网+、大数据与商业创新国际会议”论文集

冷柏军　罗　欣　主编

对外经济贸易大学出版社
中国·北京

图书在版编目（CIP）数据

2017“互联网+、大数据与商业创新国际会议”论文集 / 冷柏军，罗欣主编. —北京：对外经济贸易大学出版社，2018.11

ISBN 978-7-5663-1879-4

Ⅰ. ①2… Ⅱ. ①冷… ②罗… Ⅲ. ①互联网络–应用–企业创新–文集 Ⅳ. ①F273.1-39

中国版本图书馆 CIP 数据核字（2018）第 249901 号

2017“互联网+、大数据与商业创新国际会议”论文集

冷柏军　罗　欣　**主编**

责任编辑： 王　煜

对外经济贸易大学出版社

北京市朝阳区惠新东街 10 号　　邮政编码：100029

邮购电话：010－64492338　　发行部电话：010－64492342

网址：http://www.uibep.com　　E-mail：uibep@126.com

北京九州迅驰传媒文化有限公司印装　　新华书店经销

成品尺寸：185mm×260mm　　31.5 印张　　728 千字

2018 年 11 月北京第 1 版　　2018 年 11 月第 1 次印刷

ISBN 978-7-5663-1879-4

定价：98.00 元

编 委 会

前　言

互联网+、大数据等现代新兴理论和技术迅猛发展，已成为当代商业创新领域中重要的工具。2017 年 7 月 8—9 日，2017“互联网+、大数据与商业创新国际会议”在对外经济贸易大学举行。本次会议涉及管理科学与工程、工商管理、计算机科学、经济学及其他交叉学科等多个领域，旨在为研究者和实践者提供一个互动交流的平台，搭建一个互联网背景下 IT 创新和商业实践之间的桥梁和纽带。对外经济贸易大学校长王稼琼教授、教育部电子商务类专业教学指导委员会主任刘军教授、对外经济贸易大学信息学院院长冷柏军教授、美国新墨西哥大学罗欣教授、北京博导前程信息技术股份有限公司段建董事长等出席并致辞。

本次会议包括主题演讲、“与主编面对面”以及论文宣讲分论坛三项议程。在主题演讲阶段，会议邀请到教育部管理科学与工程专业教指委主任、清华大学陈国青教授，国际信息系统协会前任主席、哈尔滨工业大学“外专千人计划”Doug Vogel 教授，《信息和管理》（*Information & Management*）期刊主编、香港大学商学院 Patrick Chau 教授，国际顶级期刊《信息系统协会期刊》（JAIS）主编、弗吉尼亚大学 Mclntire 商学院 Suprateek Sarker 教授，《当代国际酒店管理期刊》（IJCHM）主编、美国中弗罗里达大学 Fevzi Okumus 教授，中国社会科学院经济研究所裴长洪研究员，对外经济贸易大学国际商学院王永贵院长分别进行主题演讲。“与主编面对面”环节是此次会议的特色环节，此环节由罗欣教授主持，三位国际顶级期刊主编和参会者进行面对面交流，解答参会者提出的高质量论文创意思路构建、撰写及发表等相关问题，引发热烈讨论。此外，会议还设置“大数据与商务智能”“互联网+时代的新兴问题”“电子商务与商业创新”与“管理科学与行为学研究”四个分论坛，来自世界各地的多位作者进行了研究内容的分享。本次国际会议的召开，使国内外学者同仁能共同探讨相关领域的前沿热点问题，并就各自的研究成果进行交流分享，整合优秀的学术资源，为信息科学、数据科学和商业管理的创新与未来发展增添新的活力。

本次会议面向国内外专家学者征稿，我们从征稿中选取了五十余篇学术论文汇编成册，以展示互联网+、大数据和商业创新的最新研究成果，并希望能够进一步推进我国管理科学与工程的学科建设。

本论文集由对外经济贸易大学信息学院组织论文收集和联系出版工作，并得到了对外经济贸易大学科研处和学科办的资助，在此对支持和参与编著的所有师生表示感谢；同时，我们也期盼得到学术界同行们的斧正。

编　者

2018 年 5 月

目　　录

The Effect of Data Quality on Decision-Making in Mobile Environments

—A Framework for Experimental Research

Reinhard Bernsteiner, Peter Mirski

Management Center Innsbruck, Innsbruck, Austria
Department Management, Communication & IT-Analytics Lab

Abstract: In today's fast changing and competitive business environments it is necessary to support decision makers with high-quality data. At the same time smartphones and other mobile technologies are a substantial element of our daily life. Being mobile, increases the independency and flexibility of carrying out tasks. There has been a significantly change in the way we communicate, carry out tasks and make decisions. In this paper, we develop a framework for experimental research to empirically measure data quality in the mobile environment. The specific characteristics of mobile devices may change the perception of data quality. An experimental setting is proposed to investigate and explore the effects of changing data quality on mobile devices in terms of decision-making performance and time needed to make decisions. The context of this experiment is business oriented and has its focus on the data consumer's perspective.

Keywords: data quality measurement, decision-making, experimental research, mobile device

1. Introduction

Information is the fourth production factor beside labor, capital and soil. In today's information age in some organizations data is the most important asset[①]. As for any other product data has to fulfill quality requirements.

① Giese, A. Informationsmanagement 2.0. Wegweiser für eine moderne Informationsversorgung in Unternehmen, 1st edn. Recht-Wirtschaft-Steuern, Igel-Verl., Hamburg, 2009

Kathri[①] states that data quality is very similar to product quality, as it also refers to the ability to satisfy its usage requirements. Apart from usage requirements, that can differ dependent on the end user's needs, data quality has multiple dimensions. In literature, there are several opinions regarding data quality dimensions available.

Data quality has effects on various aspects of a company, like customer satisfaction, operational costs, effectiveness of decision-making, and strategy creation and execution[②].

As shown by Brockmann et al.[③], mobile devices offer a lot of possibilities and new areas for businesses. New potentials can be seen in supporting all kinds of work tasks and to increase the efficiency of decision makers.

It can be seen that the traditional office workspace is not the most significant place to work anymore, therefore by the introduction of mobile devices employees get the independence to work from everywhere.

However, with the introduction of the mobile devices into the daily work routine also some challenges and needs further investigation of the usage patterns in different application scenarios. Depending on these scenarios the effects of data quality has to be reevaluated to meet the user's requirements.

2. Data Quality

Poor data quality can endanger the effectiveness of the strategies as well as the daily operations of an organization. As a consequence, this can lead to serious problems not only for departments but also for the entire organization[④⑤].

The economic consequences of poor data quality for decision making have been explored in several publications[⑥–⑨].

2.1 Data quality approaches

A shift from a rather technology-driven economy to an economy, where appropriate management of data serves as the key to success, is recognizable. The main goal of

① Khatri, V., Brown, C.V. Designing data governance, Commun, ACM, 2010, doi: 10.1145/1629175.1629210

② Redman, T. The impact of poor data quality on the typical enterprise, Communications of the ACM, 41(2), 79-82, 1998

③ Brockmann, T., Stieglitz, S., Kmieciak, J., Diederich, S. User acceptance of mobile business intelligence services. In: 2012 15th international conference on network-based information systems (NBiS), Melbourne, Australia, 26.09. 2012-28.09.2012, pp. 861-866. IEEE, 2012, doi: 10.1109/NBiS.2012.129

④ Redman, T. The impact of poor data quality on the typical enterprise, Communications of the ACM, 41(2), 79-82, 1998

⑤ Fisher, C.W., Kingma, B.R. Criticality of data quality as exemplified in two disasters, Information & Management, 39(2), 109-116, 2001

⑥ Fisher, C.W., Chengalur-Smith, I., Ballou, D.P. The impact of experience and time on the use of data quality information in decision making, Information Systems Research, 2003, doi: 10.1287/isre.14.2.170.16017

⑦ Pipino, L.L., Lee, Y.W., Wang, R.Y. Data quality assessment. Commun, ACM, 2002, doi: 10.1145/505248.506010

⑧ Jung, W., Olfman, L., Ryan, T., Park, Y.-T. An experimental study of the effects of contextual data quality and task complexity on decision performance. In: IRI -2005 IEEE International Conference on Information Reuse and Integration, Conf, 2005. Las Vegas, NV, USA, Aug. 15-17, pp. 149-154. IEEE, 2005, doi: 10.1109/IRI-05.2005.1506465

⑨ Batini, C., Cappiello, C., Francalanci, C., Maurino, A. Methodologies for data quality assessment and improvement. ACM Comput, Surv. 2009, doi: 10.1145/1541880.1541883

data management is to manage data as an asset of an organization, and thereby supports strategic goals and assures business success.

The quality of data is crucial and must fulfill the context—specific requirements of different stakeholders. Evolving trends such as mobile devices, cloud computing, or big data permanently create new challenges for data quality management.

The research with a focus on data quality dates back to the late 1980s at MIT. Table 1 gives a structure and comprehensive overview of the different streams in data quality research.

The research framework which is developed in this paper is situated in a business context and wants to investigate the effects of varying data quality on the performance of organizational units and individuals. The economic results of decisions are measured.

2.2 Data quality models

For this research, Wang and Strong's Quality Framework, which comprises of 16 different data quality dimensions, clustered into four categories (intrinsic, contextual, representational, and accessibility data Quality) is used.

Wang and Strong define data quality as "data that are fit for use by data consumers", and they define a data quality dimension as "a set of data quality attributes that represent a single aspect or construct of data quality"①.

This model has its focus on the view of people using data as their leading assumption is that "data consumers have a much broader data quality conceptualization than IS professionals realize"②.

This framework can be used for measuring, analyzing and improving information quality in scientifically valid ways. Utilizing it enables information systems professionals to comprehend and meet information consumers' quality requirements. Researchers argue that the framework is complete from the viewpoint of information consumers and it is methodologically sound③.

Table 2 shows the four data quality categories with their dimensions.

2.3 Research questions and methodology

Eighty-one responses from B2B customers were collected. The authors employed a mixed methods research design④ by analysing quantitative data as well as explicit comments by customers.

Intrinsic data quality: This category indicates that information possesses quality in its own right. It consists of the dimensions accuracy, objectivity, believability and reputation. The quality of the data is intrinsic

① Wang, R., Strong, D. Beyond accuracy: What data quality means to data consumers. Journal of Management Information Systems, 12(4), 5-33, 1996

② Wang, R., Strong, D. Beyond accuracy: What data quality means to data consumers. Journal of Management Information Systems, 12(4), 5-33, 1996

③ Wang, R., Strong, D. Beyond accuracy: What data quality means to data consumers. Journal of Management Information Systems, 12(4), 5-33, 1996

④ Lee, Y.W., Strong, D.M., Kahn, B.K., Wang, R.Y. AIMQ: a methodology for information quality assessment. Information & Management, 2002, doi: 10.1016/S0378-7206(02)00043-5

when the quality of the data is directly knowable of the data.

Accessibility data quality: This category describes how data is available or obtainable by the data user and points out the role of the system which provides data in terms of data operations. The system also has to be secure.

Contextual data quality: Emphasizes that information quality is dependent on the context of the task. The context primarily refers to the task of the data user (data consumer) but not the context of representation itself. This category includes the dimensions relevancy, value added, timeliness, completeness and appropriate amount of data. Data needs to be illustrated in a way by the system, so that it is interpretable, easy to understand, and represented concisely and consistently.

Representational data quality: This category reflects the importance of the presentation of data. It consists of the dimensions interpretability, ease of understanding, representational consistency and conciseness of representation. Wang and Strong[①] describe representational consistency as data that is continuously presented in the same format, consistently represented and formatted, as well as compatible with data that was presented previously

Accuracy, relevancy, timeliness, completeness and accessibility are the most important quality dimensions for information consumers. Productive users possess for their job relevant, easily accessible, accurate, complete and timely information. Broadly speaking, the main information quality problems are missing information, existing but faulty information, and genuine, existing but unusable information[②] . From the information consumers' point of view the accessibility category causes the greatest problems, followed by intrinsic and contextual problems, and finally representational issues. 43 percent of accessibility problems are result of usability short-comings[③].

3. The Mobile Environment

As a technology that is likely to drive substantial economic impact, Internet-enabled mobile devices have found their place in our daily life.

Mobile technology gets more and more important for organizations, reasons for that are the fast-changing global competition on one hand and the increase and need of a mobile workforce on the other. Being mobile increases the independency and flexibility of carrying out tasks.

Besides these benefits some disadvantages have to be taken into consideration. The major disadvantage of mobile computing using the internet is the

① Wang, R., Strong, D. Beyond accuracy: What data quality means to data consumers. Journal of Management Information Systems, 12(4), 5-33, 1996

② Oliveira, P., Rodrigues, F., Henriques, P.R. A formal definition of data quality problems. In: Naumann, F., Gertz, M., Madnick, S.E. (eds.) IQ. MIT, 2005

③ Lee, Y.W., Strong, D.M., Kahn, B.K., Wang, R.Y. AIMQ: a methodology for information quality assessment. Information & Management, 2002, doi: 10.1016/S0378-7206 (02)00043-5

low bandwidth. Wireless networks are made available by communications providers are lacking when compared to a wired or wireless network[①]. To use mobile devices efficiently, it is essential to provide a high network bandwidth between the mobile device and the relevant servers.

For mobile services, it is important that the devices are always connected to the network and servers. In the event of network failures, traffic congestions or when services are unavailable or unserviceable, mobile users will be unable to execute the desired application and commit their tasks[②].

Besides the limitation of bandwidth there are also some physical restraints associated with the device. Battery life and storage capacity are key limitations in regard to mobile applications running in the cloud. Therefore, there is an ever-increasing demand to design and to develop applications which consume less energy resources[③]. Based on limited computing capability smart devices have a much lower capacity in processing power (CPU), memory and storage when compared to the performance of laptops or personal computers[④].

The size of the screen is another important factor when it comes to usability and handling function of mobile devices[⑤]. Often it is not easy to optimize and visualize a proper graphical user interface on small screens.

Research on data quality for mobile devices usually deals with the visualization and presentation of data. Research projects are designed for laboratories and field settings as well[⑥⑦]. Presenting data visually helps the user to understand the meaning of data much faster and precise, thus it is an important research field. But the presentation of data is only one element of the data quality dimensions used for this research framework and can be found in the dimension "Representational Data Quality". When it comes to decision making in a business context all relevant data quality dimensions have to be taken into consideration. It is not

① Dinh, H.T., Lee, C., Niyato, D., Wang, P. A survey of mobile cloud computing. Architecture, applications, and approaches. Wirel. Commun, Mob. Comput, 2013, doi: 10.1002/wcm.1203

② Dinh, H.T., Lee, C., Niyato, D., Wang, P. A survey of mobile cloud computing. Architecture, applications, and approaches, Wirel. Commun. Mob. Comput. 2013, doi: 10.1002/wcm.1203

③ Gupta, P., Gupta, S. Mobile cloud computing: The future of cloud, International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, 1(3), 134-145, 2012

④ Qi, H., Gani, A. Research on mobile cloud computing. Review, trend and perspectives. In: 2012 Second International Conference on Digital Information and Communication Technology and its Applications (DICTAP). Bangkok, Thailand, 16.05.2012-18.05.2012, pp. 195-202. IEEE, 2012, doi: 10.1109/DICTAP.2012.6215350

⑤ Qi, H., Gani, A. Research on mobile cloud computing. Review, trend and perspectives. In: 2012 Second International Conference on Digital Information and Communication Technology and its Applications (DICTAP). Bangkok, Thailand, 16.05.2012-18.05.2012, pp. 195-202. IEEE, 2012, doi: 10.1109/DICTAP.2012.6215350

⑥ Chittaro, L. Visualizing Information on Mobile Devices, Computer, 2006, doi: 10.1109/MC.2006.109

⑦ Nieminen, M., Röykkee, M. (eds.) Proceedings of the 8th conference on Human-computer interaction with mobile devices and services-Mobile HCI '06. the 8th conference, Helsinki, Finland, 12.09.2006-15.09.2006. ACM Press, New York, USA, 2006

enough to explore only a single aspect of data quality, a comprehensive approach is required.

4. Framework for Experimental Research

The following section describes the research framework which is designed to collect and analyze empirical data.

4.1 Research question

The central aim of this research framework is to provide a basis that allows answering the research question, if variations in data quality in a mobile environment lead to variations in decision-making.

Based on this general research question, a) the influence on time a participant takes to make a decision as well as b) the influence on decision-making performance is to be investigated in detail. Based on this research direction hypotheses have to be extracted.

4.2 Context of the experiment

The experiment is embedded in a business context and measures the impact of data quality. The participants act as managers in a purchase department. They are responsible to order the right amount of a specific product that best fits future demands.

Their goal is to order the exact number of products needed in future periods. Ordering too many products results in storage costs, these costs can be divided in fixed storage holding costs and variable costs for each product on stock. Ordering too few products for future periods leads to additional costs generated by unsatisfied clients. No additional costs arise if the exact number of products is ordered. This situation is the most successful situation and this is the best decision performance.

4.3 Design of the experiment

The general design of the planned research project is a quantitative approach. An experimental design was selected to harvest data from the participants.

In experimental settings, "the researcher actively intervenes in the normal pattern of variation, systematically controlling variation in the independent variable (or variables) to assess the cause impact(s) of this variation on some outcomes(s)."①

Related to this statement the data quality presented to the participants is altered in order to identify and assess causal effects on decision-making. Several different groups are formed and each group has to go through the same scenario, but different levels of data quality are presented to the participants.

An application for mobile phones has to be developed with which the participants have to interact in the experiment. Using this application, the data are presented and the participants have to make their decisions.

The data are presented in different formats like raw data, graphs or tables to simulate different levels of data quality. In addition to that, some data are incorrect or incomplete.

---

① Crano, W.D., Brewer, M.B. Principles and methods of social research, 2nd edn. Lawrence Erlbaum Assoc, Mahwah, NJ, 2002

Data with different levels of data quality are presented to the different groups with the assumption that different data quality leads to differences in decision-making in terms of performance and time needed to make the decision.

4.4 Procedure of the experiment

The participants have to be selected thoroughly as they should be representatives of different age groups, educational backgrounds, occupations or gender. Based on these demographic values the participants have to be associated to the different treatment groups, which get data in different quality levels during the course of the experiment.

Before the experiment starts the participants run through the informed consent process. They are informed on the purpose of the experiment and that no personal identifiers are used and all data stored remain confidential.

After that the specific course of the experiment is explained. They have to analyze the presented data and make their decisions in their role as a manager in a purchase department and order goods.

At the very end of the online experiment, subjects were asked to provide their demographic data.

4.5 Assessing data quality

Lee et al.[①] propose AIMQ methodology to assess the quality of data. A questionnaire with 15 measures is one component of the AIMQ methodology which is used as a basis for this empirical research. All items are measured on a 0 to 10 point Likert-scale.

This methodology has been applied in various organizations in different industries and it can be used to identify IQ related problems, identify areas of data quality improvements, and monitor these improvements over time.

4.6 Preliminary model for data analysis

It is planned to calculate two dependent variables that are derived from the independent variables:

Time used to make decisions: This is measured in seconds, and is recorded in the background. Participants are not told that time was measured in the experiment, since this could have an impact on the results. The participants have to work on several decisions within one scenario. The time that is used to make decisions is calculated and stored, and then added to the total amount of time that it takes for respondents to complete one scenario.

Economic performance of the decisions: This is measured in terms of total success (e.g. profit, reduce cost, etc.) that participants are able to make in the scenario. More accurate estimations would have a positive effect on success, whereas less accurate estimations would lower the success.

All in all, the information quality

① Lee, Y.W., Strong, D.M., Kahn, B.K., Wang, R.Y. AIMQ: a methodology for information quality assessment. Information & Management, 2002, doi: 10.1016/S0378-7206(02)00043-5

model consists of 16 items whereby not all items can be tested in a valid way in the planned experimental setting. A thorough selection of the relevant and testable items has to be made. The final decision which items are used has not yet been made. Once the selection is done, it is planned to use these items as the independent variables.

On basis of the results the model will be updated and enlarged, especially the meaning of demographic variables has to be investigated. Demographic variables could be used as moderating variables.

5. Summary and Outlook

This research prepares the basis for an empirical examination and investigation of the impact of variations in data quality in decision-making in the mobile environment. The proposed research model has a clear business focus which measures the impact of data quality in terms of performance, cost or benefits or results of operations.

The next logical step is to design the experiment in detail, test and evaluate it, and run it with well-selected participants.

The results can be used to improve mobile environments for decision-making and to achieve a common understanding on data quality within an organization. Once this common understanding is created concrete goals on data quality can be defined and appropriate measures can be undertaken to reach these goals. From a data user's perspective, business processes, systems or applications in have different need in terms of data quality, which have to be identified first.

Still more research is needed to assess the impact of data quality in different industries, environments or technologies like Big Data, Cloud Computing and Internet of Things. It has to be explored how these emerging technologies influence data quality and what are the expectations on data quality from different stakeholders.

Table 1 Topics of Data Quality Research Source①

Topics
Data quality impact
Application area (e.g., CRM, KM, SCM, and ERP)
Performance, cost/benefit, and operations
IT management
Organization change and processes
Strategy and policy
Database-related technical solutions for data quality
Data integration and data warehouse
Enterprise architecture and conceptual modeling
Entity resolution, record linkage, and corporate house holding
Monitoring and cleansing
Lineage, provenance, and source tagging
Uncertainty (e.g., imprecise and fuzzy data)
Data quality in the context of computer science and IT
Measurement and assessment
Information systems
Networks
Privacy
Protocols and standards
Security
Data quality in curation

① Zhu, H., Madnick, S.E., Lee, Y.W., Wang, R. Data and information quality research: Its evolution and future. In: Topi, H., Tucker, A.B. (eds.) Information systems and information technology, Computing handbook, ed.-in-chief. Allen B. Tucker; 2, 3rd edn., 1-20. CRC Press, Boca Raton, Fla. 2014