A Training Tool for Information Quality

Markus Helfert, Mouzhi Ge

Abstract: *Over the last years information quality gained increasingly importance in practice as well as academia. Recently aspects of information quality are included in curricula. However, teaching information quality to the students is challenging and often with a theoretical focus. As a consequence many graduates have a limited understanding of information quality issues and its importance. In order to help teaching students information quality issues, we propose a teaching tool that can help the students to understand the importance of information quality management. The tool can show the effects of information quality on organizational decision-making.*

Key words: *Information Quality, Information Quality Education, Effects of Information Quality*

INTRODUCTION

Over the last decade, the importance of information quality (IQ) is ever more recognized among practitioners and academics. It has developed beyond the traditional view of IQ as a synonym for data accuracy. Wang and Strong (1996) note that in order to improve data quality in an organisation, a multi-dimensional view of the concept must be taken.

Many researchers have examined IQ. The result of the plethora of publications is a multiplicity of descriptions, definitions, criteria lists, case studies and frameworks for various areas of application (e.g. Wang, Storey & Firth 1995). In addition to these frameworks, literature on IQ indicates a number of technical, managerial, and organisation factors that are believed to improve IQ. However, at present there are few tools used for IQ training and education. Furthermore, there is no practical tool focusing on teaching IQ skills or competences. In order to provide a tool aiming to increase the awareness of IQ, we propose a teaching tool to show practical application of IQ.

This paper is organized as follows. Section 2 provides a review of IQ research to highlight the development of IQ. Followed by the review, section 3 proposes a teaching tool that can be used to show the effects of IQ in organisations. In order to evaluate the tool, section 4 presents some comments received from users using our tool. Finally, section 5 concludes this paper by summarizing our experiences and outlines the further developments.

LITERATURE OVERVIEW

Over the last decade, the importance of information quality (IQ) is ever more recognized among practitioners and academics. It has developed beyond the traditional view of IQ as a synonym for data accuracy. Wang and Strong (1996) note that in order to improve data quality in an organisation, a multi-dimensional view of the concept must be taken.

Over the last decades IQ became an important area both in research and practice. As recent job openings indicate, many organisations require professionals with expertise as IQ Manager, IQ Analyst, IQ Consultant, and IQ System Developer. Addressing the growing demands for qualified IQ professionals, taught courses and research programmes were initiated at some Universities. One example for such programmes is the IQ Program at the University of Arkansas at Little Rock (USA). Professional training courses are offered from various organisations, such as the IQ Programme at the Massachusetts Institute of Technology, Training programmes from the International Association for Information and Data Quality or Information Impact International, Inc.

Content taught within these programmes can build on an extensive foundation of IQ research. From the middle 1990's to present, IQ research becomes intensive, systematic and empirical. Therefore, the amount of IQ papers significantly increases in

a wide range of journals and conferences. From 1995 to 2008, more than 15 IQ books are published. These books have addressed different aspects of IQ research. Three IQ journals have been launched so far: Data Quality Journal in 1995, International Journal of Information Quality in 2007 and ACM Journal of Data and Information Quality in 2008. Also, many leading database and information system conferences such as SIGMOD, VLDB and CAiSE have included IQ as one of the conference themes. Furthermore, since 1996, International Conference on Information Quality (ICIQ) is annually held to provide a forum for researchers and practitioners to present research findings and exchange IQ knowledge.

From our review, we identified following main concept and key points that current IQ training programmes and courses contain or aim to achieve:

- Demonstrate a critical awareness of the importance and implications of IQ.
- The Integration into E-Learning Platforms
- Appraisal of various definitions for IQ and a thorough understanding of IQ as fitness for use in a particular application.
- A systematic understanding of concepts, principles, tools, and models essential in defining, measuring, analyzing, and improving the quality of information.
- A systematic understanding of Information Science theories and practices in the areas of database systems, systems analysis, and information visualization
- A systematic understanding of interrelationships between IQ and other key information systems concepts such as enterprise architecture, data warehousing, analytical information systems, data integration, data modelling.
- The ability to develop IQ strategies, policies, and programs to support an organization's operational, tactical, and strategic needs
- The ability to critically evaluate problems and alternative solutions in a variety of context, such as customer relationship management, logistics or Web environments.
- Awareness of ethical standards of the profession such as data privacy and protection which aims to ensure compliant use of the IQ expertise.

However, despite the increasing number of IQ training courses, only few information system professionals have received formal training or education to manage IQ (Khalil et al. 1999). Furthermore many educational programmes of IQ focus on theoretical aspects. Recognising the limitation of current efforts to teach IQ, we aimed to provide a different learning experience for students.

Our approach was developed in the context of an undergraduate "Information System Strategy" Course using a problem-based teaching approach. As foundation we use the textbook from Chaffey and Wood (2005), which provides a problem and case study based approach to information systems. In order to emphasise the concepts taught we include case studies and discussions throughout the course. Furthermore, in order to provide a practical experience for students, we developed a software tool in form of a game to help students understand the importance of IQ management.

TOOL DESCRIPTION

The tool is based on the traditional Beer Game, which involves managing supply and demand in beer supply chain. The concept for this game was first developed at the Massachusetts Institute of Technology in 1960s. Since then, several extensions and modifications are proposed. Kaminsky and Simchi-Levi (1998) identified several weaknesses of this traditional game and extended to the computerized Beer Game.

We extended the traditional Beer Game and included IQ aspects in form of various marketing and sales information to the students. The system is designed, to information of different quality levels. Using the given information, students are asked to make inventory control decisions.

This supply chain game involves manufacturer, distributor and customer. The participants are asked to play the distributor and the other two roles are taken over by the computer. According to the inventory information and customer ordering history, participants will order products from the manufacturer and supply products to the customer. There are two sources of cost associated with the game: (1) If the distributor cannot fill customer's order, a cost will be occurred (1 euro per unfilled item). (2)When the items are stored in the inventory, a cost will be occurred (0.5 euro per stored item per week). The goal of the game is to minimize the cost in the distributor's inventory management.

The interface of our teaching tool is shown in Figure 1. In order to start the game, participants place an order in the input box and click the "OK" button. The order is sent to the manufacturer. The procurement delay from manufacturer takes 3 weeks. That means it takes 3 weeks from making the order and receiving the ordered items. Note that the manufacturer may not be able to fill the order if this order exceeds the inventory of the manufacturer. For example, in week 10, participants place an order to the manufacturer. This order will arrive to the manufacturer in the beginning of week 11. The manufacturer needs 1 week to process the order. In the beginning of week 12, the manufacturer ships the ordered items to the distributor and the shipment takes 1 week. Therefore if participants make an order in week 10, the order will arrive in week 13.

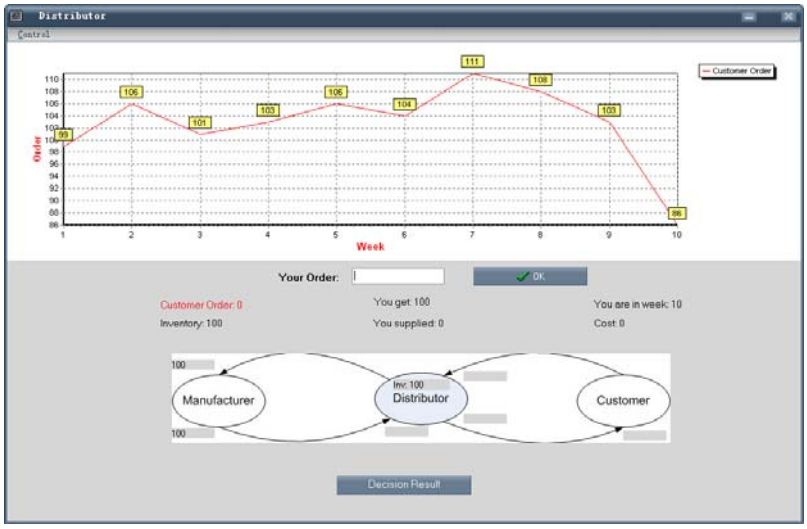

Fig.1. Initial interface of this teaching tool

When playing the game, participants could obtain the following information from the software interface: customer order, current inventory, how many ordered items are arrived, how many items are shipped to the customer, current week, and the total cost. Participants also could observe the delays between the manufacturer, the distributor and the customer. One round of the game contains 10 weeks. Figure 2 has shown snapshot of the software after one round of the game. After completing 10 weeks, the student can play another 10 weeks or finish the game. When finishing the game, the student is shown the results.

When participants click the decision result button, the decision results of last 10 weeks are showed as Figure 3.

As indicated in Figure 3, students can observe the relationship between IQ and decision quality. For instance, the standard deviation of information accuracy is 10 and

the information completeness is 80%, the participant made 10 ordering decisions and the total cost is 356 euro.

The measurements of completeness and accuracy are organised as follows. The levels of completeness are expressed by percentages. The percentages can be determined by the equation 4-5. In the experiment, 5 completeness levels are used: 20%, 40%, 60%, 80% and 100%. For example, 60% completeness means only 60% of the information is provided.

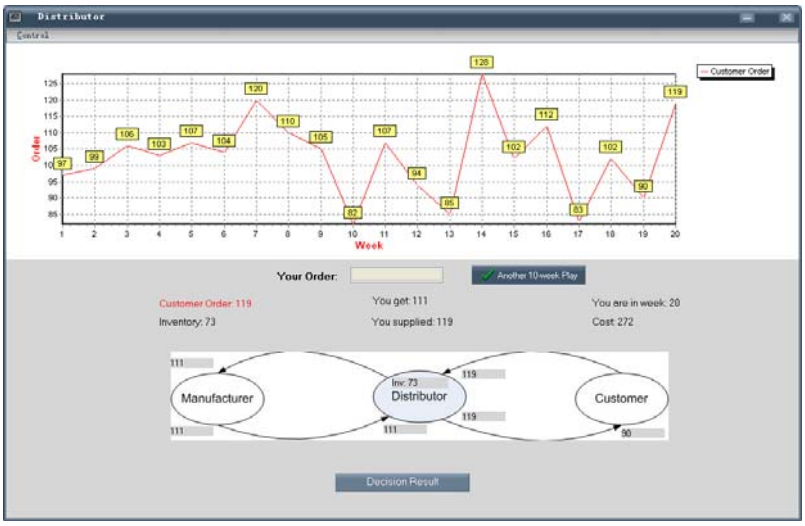

Fig.2. Snapshot of one round of the game

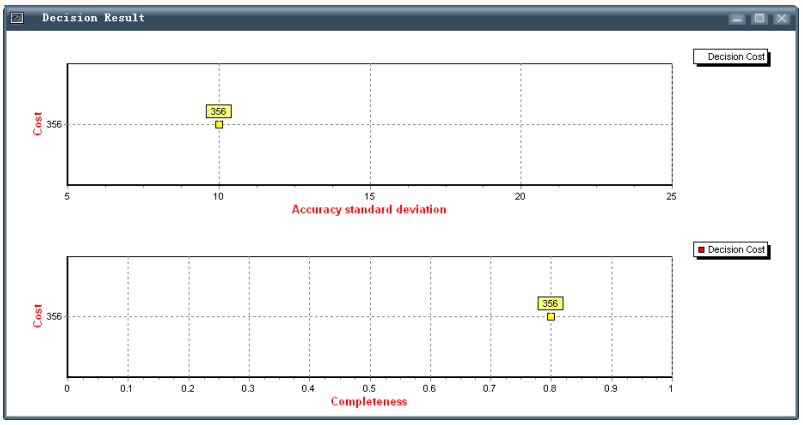

Fig.3. Decision results

Accuracy is divided into 5 levels by different standard deviations. These 5 standard deviation levels are 5, 10, 15, 20, and 25. The greater the standard deviation, the greater the degree of inaccuracy is. For example, the orders generated by standard deviation 20 is always more inaccurate than the ones generated by standard deviation 5. The reason why we employ the standard deviation is that we can design the inaccuracy distributed in every order.

TOOL EVALUATION AND FURTHER IMPROVMENTS

The tool was developed in the context of a taught undergraduate course and used with a blended learning approach (Traditional Lecturers supported by E-Learning). At present, we received and incorporated feedback from students; however the tool was yet not formally incorporate into the course. Furthermore a complete integration into our E-Learning Platform was not realised yet. This is envisaged for the next academic year. In the meantime, in order to receive further feedback and comments on how to improve our teaching tool, we invited 30 academics and professionals to comment on the tool. The participants are from Dublin City University, University of Oxford, Singapore

Nanyang Technological University, University of Dundee, Microsoft Research Asian, Hibernia Atlantic Ltd., Avaya Ireland, and J.P. Morgan UK. 65% of subjects were male and the other 35% are female. The average age of subjects is 32.

All the participants' works or studies are related to information processing. This information can be presented by text, audio, video, conversation etc. The most common information processing concerned checking emails and collecting business information. Some participants' work concluded with a report. This report can be considered as a result of gathering and processing different information. Therefore IQ is directly related to the quality of their work. Some participants' work is centred on data analysis and information management. IQ is vital since poor quality information may generate erroneous analysis results, which could incur a variety of business losses.

After the tool evaluation, we carried out a semi-structured interview to the participants. This interview consists of two major themes: experience of using the tool and opinions towards the IQ problems. After collating the answers, we organise and interpret the feedbacks as follows.

Through using this tool, the interviewees stated that this tool is a practical demonstration to show the effects of IQ. Compared to theoretical training, the tool increases the interest and understanding to IQ concepts. It also helps the participants to raise the awareness of IQ management. Many interviewees emphasized that IQ is observed as a crucial factor to the success of information systems. Organisations need to cultivate IQ concepts into organisational cultures.

Additionally, the interviewees found the tool is easy to use and can help non-IQ-professionals to quickly understand IQ concepts. For example, they realize that IQ can be measured by different dimensions such as accuracy and completeness. Importantly, after using the tool, the participants understood how information accuracy and completeness perform in a practical application. Therefore the interviewees concluded that improving IQ could enhance their work performance. Organizations need to increase the awareness of IQ management.

When concerned with IQ problems, Most of the interviewees experienced IQ problems in their work or daily life. Interviewees provided us with various examples of such IQ problems. Compiling their feedbacks, we found that IQ problem is highly pervasive in work and everyday life. Hence it is valuable to improve IQ in people's work and daily life. However, although all the interviewees have met different kinds of IQ problems, only a few people had considered IQ improvement. Some interviewees have grown accustomed to IQ problems and were not aware of the possibility of improving IQ. Some student interviewees complained that IQ problems always happened within their life and study, especially with regards to information from the internet. Some industrial interviewees had considered improving IQ in their company. However due to unknown budgets and the inexistence of mature IQ management models, high management always denied proposals for systematic IQ improvement. The collected feedbacks showed that although many IQ problems exist, the operation of IQ improvement procedures is very rare.

In order to find the root causes of the IQ problems, we collect the possible reasons of triggering an IQ problem. Except some typical causes such as typing errors or delayed input, interviewees are mainly concerned with two causes: system design and information processing. Some interviewees emphasized that it is important to prevent IQ problems when designing the system. For example, the lack of constraint checking in the system can result in poor-quality data in the database. Some interviewees considered that IQ problems can be generated in the information processing or transferring procedure. Therefore it is critical to increase IQ awareness when we process or transfer the information.

Besides the feedbacks above, we also collected the suggestions for the improvement of this tool. The further improvements for the tool can focus on the group

competition. That means each component of the supply chain can be played by one participant. The participants in the same supply chain are composed as a group. Thus different groups are able to compete for their performance by comparing the benefit in a determined period. In addition, we can extend the tool by including more IQ dimensions. It would enable students to observe how different IQ dimensions perform in a practical application.

CONCLUSION

In this paper, we describe a teaching tool to facilitate IQ education. The tool is based on the traditional Beer Game, originally developed at the Massachusetts Institute of Technology. We adapted this game and introduced IQ elements to the game. The tool was developed in the context of an undergraduate course in Information Systems Strategy. The course is designed as a problem-oriented course. In order to receive feedback from professionals and academics, we invited 30 subjects to use the tool and provide feedback. By using our tool the participants showed an increased awareness of IQ and its importance. In addition, they have understood that IQ is multi-dimensional concept. Although the initial feedback shows the benefit of using our tool, it is an ongoing project. We aim to improve and extend its functionality and formally incorporate the tool in our undergraduate course.

REFERENCES

[1] Batini, C. and Scannapieco, M., Data Quality, Concepts, Methodologies and Techniques, Publisher: Springer, 2006.

[2] Brodie, M. L. , Data quality in information systems. Information and Management, 3(6), 1980, pp. 245-258.

[3] Chaffey, D., Wood, D. Business Information Management: Improving Performance using Information Systems, Prentice Hall, 2005.

[4] DeLone, W.H. and McLean, E.R. , Information system success: the quest for dependent variables, Information System Research, 3(1), 1992, pp. 60-96.

[5] Kaminsky, P. and Simchi-Levi, D., A new computerized beer game: a tool for teaching the value of integrated supply chain management. In Supply Chain and Technology Management, Hau Lee and Shu Ming Ng, eds., The Production and Operations Management Society: Miami, Florida, 1998.

[6] Khalil O.E.M., Strong D.M., Kahn B.K., Pipino L.L., Teaching information quality in information systems undergraduate education, Informing Science, 2(3), 1999, pp 53-59.

[7] Paradice, D.B. and Fuerst, W.L., An MIS data quality methodology based on optimal error detection, Journal of Information Systems, 5(1), 1991, pp. 48-66.

[8] Wang, R.Y. and Strong, D. M. , Beyond accuracy: what data quality means to data consumers. Journal of Management Information Systems, 12(4), 1996, pp. 5-34..

[9] Wang, R.Y., Storey V.C. and Firth, C. P., A framework for analysis of data quality research, IEEE Transactions on Knowledge and Data Engineering, 7(4), 1995, pp.623-640.

[10] Yu, S., and Neter, J., A stochastic model of the internal control system, Journal of Accounting Research, 1(3), 1973, pp. 273-295

ABOUT THE AUTHORS

Markus Helfert, PhD, School of Computing, Dublin City University, Phone: +353 - 1700 - 8727, E-mail: Markus.Helfert@computing.dcu.ie

Mouzhi Ge, School of Computing, Dublin City University, Phone: +353 -1700 - 8980, E-mail: mge@computing.dcu.ie