2011

Enhancing Business Intelligence Traceability through an Integrated Metadata Framework

Chin-Hoong Chee
Universiti Tunku Abdul Rahman, timothychee@hotmail.com

William Yeoh
Deakin University, william.yeoh@deakin.edu.au

Shijia Gao
Monash University, caddie.gao@monash.edu.au

Enhancing Business Intelligence Traceability through an Integrated Metadata Framework

Chin-Hoong Chee
Department of Information Systems
University Tunku Abdul Rahman
Perak, Malaysia
Email: cheech@utar.edu.my

William Yeoh
School of Information Systems
Deakin University
Burwood, Australia
Email: william.yeoh@deakin.edu.au

Shijia Gao
Faculty of Information Technology
Monash University
Caulfield, Australia
Email: caddie.gao@monash.edu.au

Abstract

The Business Intelligence (BI) system provides users with multi-dimensional information (so-called BI product) to support their decision-making. However, very often business users still could not fully understand the BI product, nor have a clear picture of the entire information manufacturing chain of the BI product. In response to this situation, this paper presents an integrated metadata framework ("BIP-Map") to facilitate the traceability and accountability of a BI product following the design science research approach. Specifically, the salient modelling and management techniques from the business process modelling notation (BPMN), the information product map (IP-Map), and the metadata management are adapted to construct a three-layered integrated metadata framework enabling the business users to make timely and informed decisions. A BIP-Map informed prototype system has been developed in collaboration with online job recruitment firms. The authors conducted in-depth interviews with seven key BI stakeholders of the recruitment firms to evaluate the usefulness of the BIP-Map. It is envisaged that the metadata framework allows the technical personnel to understand the business processes that relate to certain information provided in the BI reports. Business users will also be able to gain insights into the logic behind any BI report.

Keywords

Business Intelligence, Metadata, Business Process Modeling Notation (BPMN), Information Product Map (IP-Map), Design Science

INTRODUCTION

In recent years, business intelligence (BI) application has emerged as one of the top ten spending priorities for many CIOs (Gartner 2011a). It is reported the worldwide business intelligence and analytics software revenue surpassed US$10 billion dollars in year 2010 alone (Gartner 2011b). While the BI market appears vibrant, nevertheless the implementation of a BI system is a financially large and complex undertaking, involving many different stakeholders over a lengthy period (Wixom and Watson 2010; Yeoh and Koronios 2010). According to Wixom and Watson (2010), business intelligence is "*a broad category of technologies, applications, and processes used for gathering, storing, accessing, and analysing data to help its users make better decisions*". In other words, BI is a system that allows business users to leverage the disparate data sources for making informed business decisions (Lawton 2006). However, allowing the end users to handle sets of data from different sources

creates a problem because users of respective functional units are not familiar with those data. On one hand, a BI system enables the delivery of BI products[1] from various sources to the end users (Fisher et al. 2006). On the other hand, individual business users are still facing problems in understanding the BI environment and the information provided by BI applications. They cannot visualise the entire manufacturing process of any BI product, nor understand the actual meaning of the integrated BI product that is available to them. Moreover, the complexity of the BI environment increases constantly with the number and the diversity of BI applications. So the problem of misunderstanding the information within a BI environment is getting even more challenging. Therefore, the provision of comprehensive metadata to enhance the understanding of BI products is a critical issue in the BI environment (Ballard et al. 2006; Ponniah 2001; Poe et al. 1997; Shankaranarayanan and Even 2006; Staudt et al. 1999; Vaduva et al. 2002). In fact, Gartner Research contended that metadata management is one of the most important functionalities that a BI environment should deliver (Aguilar-Saven 2003).

While prior studies investigated the technical metadata, and the metadata for the business side, there has been a relative dearth of research on the integrated metadata for the entire BI product manufacturing process. In response to this, this paper reported a design science research (Hevner and Chatterjee 2010) that develops and evaluates a comprehensive metadata framework which serves as a mechanism that provides the context about the information (i.e. product) of a BI application (Fisher et al. 2006). The metadata framework addresses the how, when, why and what questions in a BI environment (Viaene 2008). Specifically, this research has sought to integrate the modelling and management techniques used in the Business Process Modelling Notation (BPMN), the Information Product Map (IP-Map) of the Total Data Quality Management, and the Metadata Management to construct a three-layered metadata framework.

The remainder of this paper is organised as follows. Section Two presents the research background, whilst the following section introduces the integrated metadata framework (known as "BIP-Map"). Section Four illustrates the framework within the online job recruitment firms setting, followed by the evaluation findings of interviews with the seven key BI stakeholders of online recruitment firms. The final section provides a summary of the study reflecting the design science approach, and indicates future research for this study.

RESEARCH BACKGROUND

Due to today's highly competitive and fast pace business environment, contemporary managers need to make the most optimised decisions in the shortest possible time (Delen et al. 2006). Therefore, the challenge of BI systems is to extract relevant information from the huge amount of corporate data and make them readily accessible to decision makers. In this regard, metadata, which can be defined as 'data about the data', plays a crucial role in the implementation of effective BI environment (Sen 2002). Typically, metadata can be classified as business metadata (which relates to the data that is meaningful to business users) or technical metadata (which is used by information technology staffs who are responsible for developing and administering the BI system) (Ballard et al. 2006; Ponniah 2001; Poe et al. 1997; Shankaranarayanan and Even 2006; Staudt et al. 1999; Vaduva et al. 2002). That is, from business perspective, metadata is important because it provides details and explanation for a BI product and the associated data (Ponniah 2001). Whereas in the technical domain, metadata is crucial for building the data warehouse as developers need to know the data structures, source-to-target mappings, and data transformation rules during the data extraction, transformation and loading (ETL) processes (Ponniah 2001). As depicted in Figure 1 below, this general classification of metadata leads to two main types of users in the business intelligence environment, namely technical users and business users. Therefore, in view of the important role of metadata it should be compiled with both the business and technical sides in mind (Shankaranarayanan and Even 2006).

[1] Based on information product concept, a BI product refers to any reporting information that is provided to the end users for decision-making where information is treated much like the products found in a manufacturing environment of any industry (Wang 1998; Wang et al. 1998; Shankaranarayanan et al. 2000; Shankaranarayanan 2005).

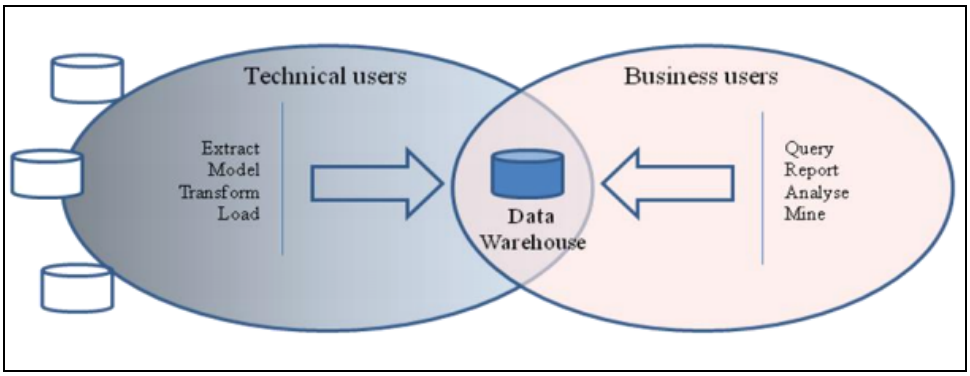

Figure 1: General Users in a Business Intelligence Environment

Although metadata has been in place for most BI systems, very often the end users still face problems in understanding the relationship of all participants and processes for a business transaction (Tan et al. 2008). The business process of an organisation is "*a set of activities that are performed in coordination to realise a business goal*" (Weske 2007). To represent the processes of an enterprise, the Business Process Modelling Notation (BPMN) is the technique used so that the current processes may be analysed and improved in the future. It is important to identify clearly the activities and participants that are involved in the business process since a BI product can only be analysed if its business processes are being modelled correctly (Aguilar-Saven 2003). The BPMN is a very useful modelling technique for managing the business processes as it presents to the users a clear workflow of the transactions involved in the business. The BPMN, however, does not describe the manufacturing process of data and so users will not understand how corporate data is captured, processed, stored and utilised. This problem can be addressed, in part, by the use of an Information Product Map (IP-Map) which describes the manufacturing processes of data in a detailed manner.

Originating from the Total Data Quality Management (TDQM) concept, the purpose of IP-Map is to deliver high quality information to users by treating data and information in much the same way as other products used in the manufacturing industry (Wang 1998; Wang et al. 1998; Shankaranarayanan et al. 2000; Shankaranarayanan 2005). It allows people to comprehend, evaluate, and describe how information is compiled. Yet, an IP-Map still does not provide descriptions of the business processes that are related to the manufacturing or utilisation of data in an organisation and so users are unable to identify the business processes that have an impact on corporate data. Therefore, the adoption of BPM and IP-Map modelling techniques will complement each other.

BUSINESS INTELLIGENCE PRODUCT AND METADATA MAP (BIP-MAP)

Based on the modelling and management techniques used in Business Process Modelling Notation (BPMN), Information Product Map (IP-Map), and metadata, this section presents the conceptual metadata framework, known as Business Intelligence Product and Metadata Map (BIP-Map). The purpose of this integration is to produce a comprehensive metadata framework which is able to enhance the traceability and accountability of BI so that the end users can make informed decisions for their business.

The BIP-Map is built as a multi-layered framework where the business processes, information processes, and metadata are grouped into three different layers, with BPMN as the first layer, IP-Map as the second, and metadata as the third. Figure 2 depicts the high-level conceptual framework of BIP-Map.

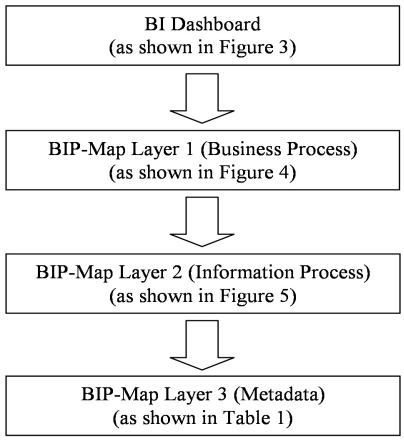

Figure 2: The High-level Architecture of BIP-Map

BIP-Map Layer 1 (business process layer) presents the workflow of business processes to the users so that they can understand the relationship of all the participants and processes for a business transaction. In this research, a well-supported standard modelling notation, Business Process Modelling Notation (BPMN), has been applied accordingly. BPMN has been unanimously endorsed as the de-facto standard notation by the BPM Initiative Notation Working Group, which represents a large segment of the business process modelling community. After reviewing various notations or methodologies, the BPMI Notation Working Group consolidates the best ideas from the divergent notations into a single standard notation - BPMN. Based on the flow-charting techniques, BPMN provides a standardised language for business and technical developers (Wang et al. 2006; White 2004). Thus, BPMN has been integrated into the BIP-Map framework. Furthermore, when the activities and participants of a specific business process are defined at this layer, various manual tasks can be automated by the system and thus accomplished in a more efficient manner. It also has the added benefit of facilitating rapid changes and improvements to workflow and business processes. This can be achieved because the interaction of activities between different participants for a business process is indicated in this layer.

Applying the Information Product Map (IP-Map) concept, BIP-Map Layer 2 (information process layer) illustrates the information manufacturing processes so that users can identify how the data are being captured, processed, stored and utilised in order to produce better information products. Additionally, the map enables the end users to visualise the manufacture of an information product (Wang 1998; Wang et al. 1998; Shankaranarayanan et al. 2000; Shankaranarayanan 2005), to appraise its quality, and to evaluate its dimensions in order to implement continuous improvement (Fisher et al. 2006).

BIP-Map Layer 3 (metadata layer) provides metadata to users so that they can understand the real meaning of the data before using it to make important decision for their organisations. According to the work of Ballard et al. (2006), Poe et al. (1997), Staudt et al. (1999) and Vaduva et al. (2002), the metadata are grouped into a few categories to describe the entire information processes, from data entry, ETL stage till data warehousing development stage. They include metadata for data filtering/validation, metadata for data transformation, metadata for database, and metadata for data warehouse. By referring to the third layer, users are able to gain a deeper understanding for each building block in the manufacturing process. At this layer, end users can know in detail how data are being validated by the data-filtering process when they are captured at the source systems. Knowing how data are validated is very important because it can ensure that only data that are accurate and useful to the system are retained. Moreover, this layer is a useful reference for the technical people, such as database administrators and software developers, since metadata are provided for the transformation of data from a database into a data warehouse.

BIP-MAP APPLICATION: ONLINE JOB RECRUITMENT FIRMS

To illustrate the proposed BIP-Map, a BIP-Map informed prototype system was developed in collaboration with two online job recruitment firms. First, the system requirements for the executive BI Dashboard were gathered through a series of interviews with the key BI users. Next, the BI Dashboard was developed using PHP programming language and open source BI software, so-called Business Intelligence and Reporting Tools (BIRT). Both the PHP and the BIRT technologies were used to query the data from the firm's Microsoft SQL Server database into the BI Dashboard. The BIP-Map was attached to the BI Dashboard where the users can click on the link located below any graph or chart available in the BI Dashboard. The following sub-sections illustrate the three-layered BIP-Map framework based on a BI Dashboard of one the participating firms.

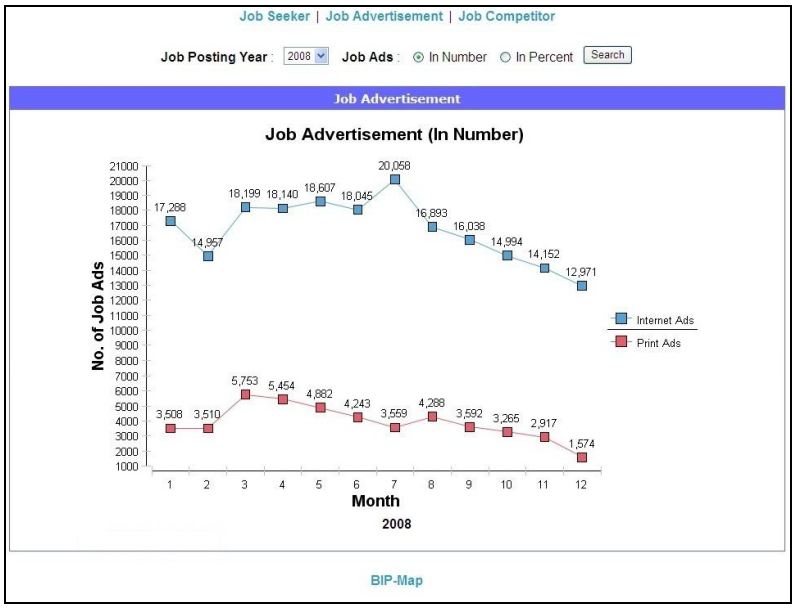

Figure 3: Job Advertisement's Graph from BI Dashboard

First Layer of BIP-Map

Figure 3 shows an example of a graph available in the BI Dashboard that displays the total number of printed and online advertisements. When the users click on the BIP-Map link located below the graph or chart on the BI Dashboard, the first layer of the BIP-Map for that particular information will be invoked. This first layer is based on BPMI Notation Working Group's BPMN, and is used to describe all the business processes that are related to the BI product. By referring to this first layer, the end users can readily identify the activities and participants of a business process relating to the manufacturing of an information product.

Figure 4 shows the first layer of a BIP-Map for the job advertisement graph that indicates how employers and customer service agents are involved with the creation of a job advertisement in the online recruitment firm. The activity represented by a dashed line is expandable, and users can click on it to access the second layer of the map for that particular activity.

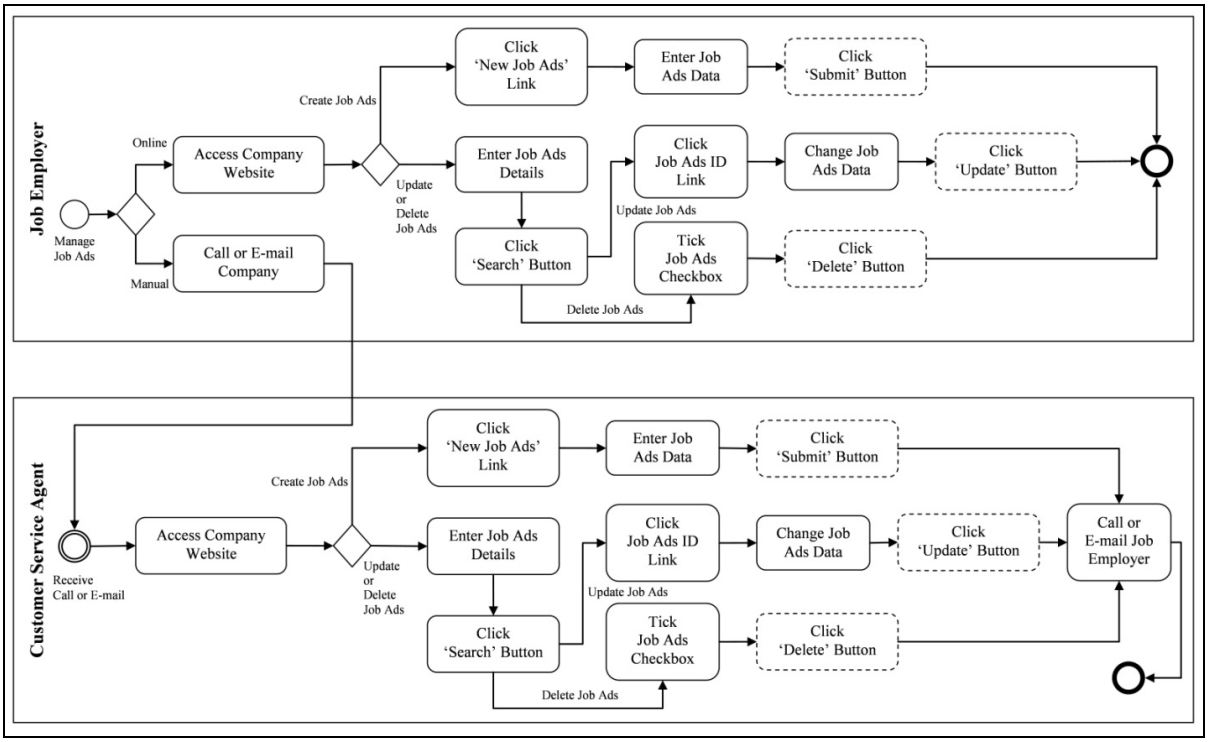

Figure 4: First Layer of BIP-Map for Job Advertisement Graph

Second Layer of BIP-Map

When the users click on any expandable process on the first layer of a BIP-Map the second layer will then be invoked. This second layer is based on the Information Product Map (IP-Map) of Shankaranarayanan et al. (2000). It is used to describe the processes that are related to the information product and users are able to visualise the steps involved in capturing, processing, and storing that product (Shankaranarayanan et al. 2000; Shankaranarayanan 2005). This helps the end users to make good use of the data in their organisations for business analysis and decision-making. Similarly, it facilitates changes to improve the workflow of information processes when the users are able to identify how the data is being captured, processed, stored and utilised throughout the organisation. Figure 5, below, shows the second layer of a BIP-Map for the job advertisement graph which indicates the manufacturing processes of job advertisement data throughout the company.

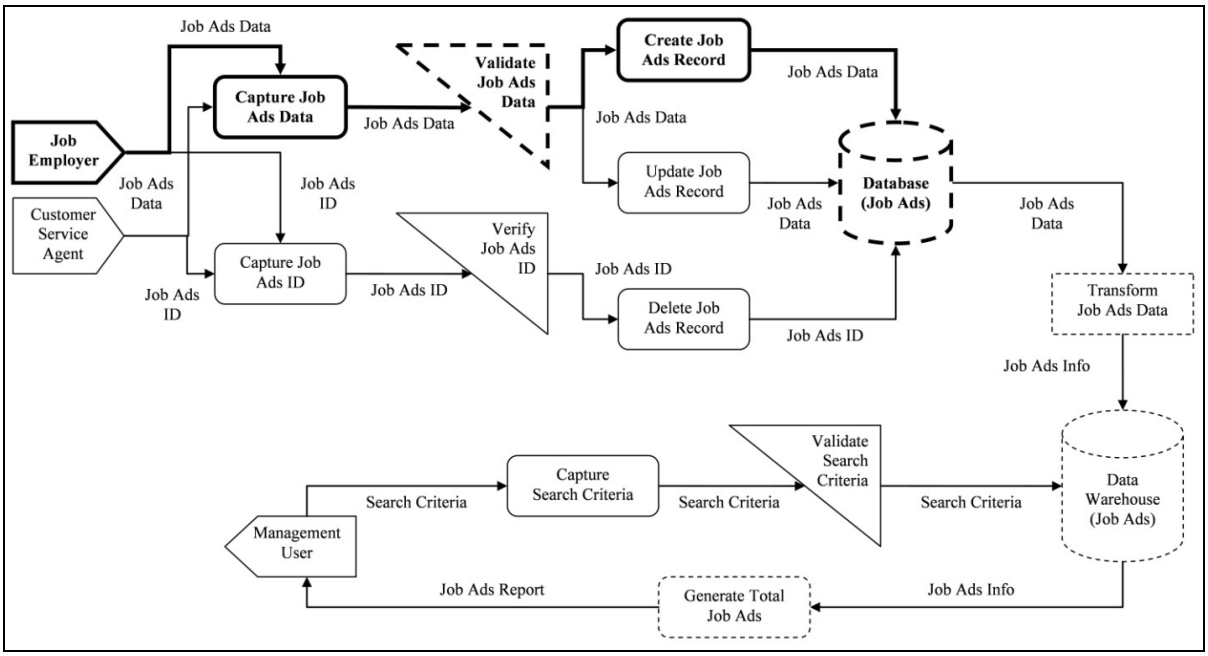

Figure 5: Second Layer of BIP-Map for Job Advertisement Graph

Third Layer of BIP-Map

Clicking on one of the dash-line building blocks on the second layer of a BIP-Map will invoke a third-layer map, which is based on the concept of metadata, and it comprises a set of metadata used to describe the processes and/or participants that are involved with the manufacturing of the information product. Table 1 represents the third layer of a BIP-Map for the job advertisement graph that shows the metadata for the data validation stage. Moreover, metadata for the ETL stage, and metadata for the data warehouse are provided in the Appendix. Based on these metadata, managers are able to identify the criteria used to generate a report from the BI system. Knowing such criteria will help them make better decisions for their organisations. In addition, metadata on the storage of data in databases and data warehouses are also a useful reference for the technical people to improve the corporate data governance.

Table 1. Third Layer of BIP-Map for Job Advertisement Graph (Data Validation)

Data	Description	Condition	Compulsory	Option
Job Position	− A name that describes the job.	− Not more than 25 alphanumeric characters.	Yes	-
Country	− A country for the job applicant to work.	− Select one option only.		− A list of countries.
State	− A state for the job applicant to work.			− A list of states for the selected country.
City	− A city for the job applicant to work.			− A list of cities for the selected state.
Job Responsibilities	− Information that describes what the job applicant needs to do.	− Not more than 500 alphanumeric characters.	No	-
Job Requirements	− Academic qualification and working experience that the job applicant needs to fulfill.			-
Application Deadline	− The job is closed for application on this date.	− A date that is greater than the current date. − The default selected date is 1 month later than the current date.	Yes	− A javascript calendar with date selection.
Advertisement Status	− Indicates whether the job advertisement is posted online or not. − Job advertisement that is saved but not posted online will not be displayed.	− Select one option only. − Default Value : Post Online		− Post Online − On Hold

EVALUATION

This research proposes an integrated metadata framework (i.e. BIP-Map) and illustrates the framework within the online job recruitment firms setting. To evaluate the usefulness of the BIP-Map, a number of interviews have been conducted with the firms' BI key stakeholders. Semi-structured interviews were selected as the primary source of evidence for this study to facilitate an examination of the organisation's experiences in relation to the integrated metadata framework (i.e. BIP-Map). Moreover, this particular technique would provide the researchers with a deeper understanding of the contextual issues than would have been possible by a more rigid form of survey (Lincoln et al. 1985; Yin 2003). It allows for a dynamic exploration and clarification of comments made by the BI stakeholders in pursuing the issues of particular significance that relate to this research. This approach allows flexibility in the sequencing of questions and in the depth of exploration. The interviewees were drawn from different functional areas which include the *Chief Executive Officer, the Chief Operation Officer, the Chief Information Officer, the Chief Technology Officer, the Regional Process Improvement Manager, the Quality and Administration Manager, and the Management Information Systems Manager*. In doing so, a standardised set of interview questions was used in all the sessions.

Face-to-face interviews with BI stakeholders were conducted on scheduled times in their respective work sites. The interviews were audio-recorded and each lasted for about an hour. During the interview, the researchers introduced the metadata framework and demonstrated the working prototype of the BIP-Map. The interviewees were then asked to comment on the artefacts based on a predefined list of questions. Hence it allows for some latitude in getting detailed information from the interviewees during the interview, but at the same time to ask general questions so that all participants are treated in a consistent manner. The data from the interview were then transcribed and analysed through a thematic analysis approach (Patton 2002).

Overall, the research participants indicated that BIP-Map is useful for business decision-making when it is being integrated into a BI Dashboard. This is because BIP-Map provides the end users with a clear picture of the information manufacturing processes and also an understanding of how the data is being tabulated into the various BI reports. Moreover, most participants acknowledged that BIP-Map enables users to gain more insight to the displayed reports or charts in a BI Dashboard. One interviewee stated, "*It helps in the understanding of how the information came about and provides a mechanism for improving the data capturing and information generation processes.*" In other words, having such a three-layered integrated metadata framework would allow the users to readily identify the data quality problems and the personnel responsible for each information product. One participant pointed out that a BI Dashboard alone is not sufficient for users to fulfil the BI needs of their organisation. Therefore, the BIP-Map will serve as a metadata framework to provide detailed information that is important for business analysis and decision-making.

In addition, the research participants maintained that each end user may have only a limited knowledge of the information manufacturing process that underlies a particular chart produced in the BI Dashboard. One interviewee mentioned that, "*This provides the technical team with some form of documentation of the report structure which can be controlled in order to ensure the report's integrity. It is also useful when the technical personnel unfamiliar with the reports need to understand the logic behind the data. With the combination of the*

three layers, the technical personnel are able to understand the business considerations behind certain data while the operations personnel can get an insight into the logic behind the report information." This indicates that users do not really understand how the data in any report are being constructed from within the firm's systems, yet when making business decisions it is risky to rely on data whose source has not been clearly identified. All the seven key BI stakeholders asserted that the BIP-Map is useful as it provides them with a clear workflow of the business and information processes together with the metadata which help them to understand the real meaning of data. It was apparent from the evaluation responses that a BIP-Map is a valuable metadata framework that can assist users to understand the entire manufacturing process of any BI product before they make any important decisions for their organisation.

SUMMARY AND FUTURE WORK

This study follows the design science approach and comprises the developing and evaluation of artefacts (Hevner and Chatterjee 2010). In this final section, we describe the research approach used and show how it follows the seven guidelines for the design science research proposed by Hevner et al. (2004).

This study involved the development of two artefacts (Hevner et al. Guideline 1 – design as an artefact): (1) an integrated metadata framework (i.e., a model) – Business Intelligence Product and Metadata Map (BIP-Map); and (2) a BIP-Map informed prototype system (i.e., an instantiation). The purpose of this research is to address the problem of users who do not fully understand the BI product manufacturing process and the actual meaning of the BI product that is available to them, and to provide a better support for the end users in their decision-making (Hevner et al. Guideline 2 – problem relevance). In achieving this, the two designed artefacts are strictly informed by, and incorporated with, three modelling and management techniques; namely the Business Process Modelling Notation (BPMN), the Information Product Map (IP-Map) and the Metadata Management (Hevner et al. Guideline 5 – research rigour). The BIP-Map comprises three layers: the first layer adopts the BPMN model to describe the business process, the second layer is an IP-Map showing information process and is constructed by expanding the first layer, whilst the third layer represents the metadata detail of data validation stage through to ETL, and data warehousing stage.

In addition, the usefulness of the artefacts are evaluated via in-depth interviews by following the interview methodology rigorously (Lincoln et al. 1985), (Yin 2003), (Patton 2002) (Hevner et al. Guideline 3 – design evaluation; Guideline 5 – research rigour). The results of the evaluation will be used to refine the designed artefacts for future works. Such a build-and-evaluate loop will iterate several times before the final design artefacts are generated (Hevner et al. Guideline 6 – design as a search process). This paper includes both the technical presentation and practical framing in terms of application in the online job recruitment domain (Hevner et al. Guideline 7 – communication of research).

The designed artefacts are the major contribution of this research (Hevner et al. Guideline 4 – research contributions). The integration of the layers into one framework enables users to quickly and efficiently understand the BI products and to be more certain in their decision-making. By integrating the BIP-Map into the BI Dashboard of an organisation, the managers will be better equipped to visualise all business and information processes relating to the manufacturing of the BI products provided to them. Moreover, metadata available in the framework will help BI stakeholders to understand the real meaning of the data and the information manufacturing process which are necessary to support their business decision. Therefore, the BIP-Map could assist BI stakeholders to easily identify the critical phases that may create data quality problems and to implement continuous improvement efforts for the BI needs of their organisations. Furthermore, the technical personnel will be able to understand the business considerations that underlie certain data while the business personnel can obtain insights into the logic used in the BI report information. In the next phase of the research, a graphical user interface (GUI-based) software will be developed to enable the rapid development of the BIP-Map.

REFERENCES

Aguilar-Saven, R.S. "Business Process Modelling: Review and Framework," *International Journal of Production Economics* (90) 2003, pp 129-149.

Ballard, C., Farrell, D.M., Gupta, A., Mazuela, C., and Vohnik, S. *Dimensional Modelling: In a Business Intelligence Environment,* IBM, New York, United States, 2006.

Delen, D., and Pratt, D.B. "An Integrated and Intelligent DSS for Manufacturing Systems," *Expert Systems with Applications* (30:2) 2006, pp 325-336.

Fisher, C., Lauria, E., Chengalur-Smith, S., and Wang, R. *Introduction to Information Quality* M.I.T. Information Quality Program, USA, 2006.

Gartner Press Release "Gartner Executive Programs Worldwide Survey of More Than 2,000 CIOs Identifies Cloud Computing as Top Technology Priority for CIOs in 2011", Retrieved 17 June 2011a, from http://www.gartner.com/it/page.jsp?id=1526414

Gartner Press Release "Gartner Says Worldwide Business Intelligence, Analytics and Performance Management Software Market Surpassed the $10 Billion Mark in 2010", Retrieved 18 June 2011b, from http://www.gartner.com/it/page.jsp?id=1642714

Hevner, A., and Chatterjee, S. *Design research in information systems: Theory and practice*. Springer, 2010.

Hevner, A.R., March, S.T., Park, J., and Ram, S. "Design science in information systems research," *MIS Quarterly* (28:1) 2004, pp 75-105.

Lawton, G. "Making Business Intelligence More Useful," *IEEE Xplore Computer Journal* (39:9) 2006, pp 14-16.

Lincoln, Y.S., and Guba, E.G. *Naturalistic Inquiry* Sage Publications, Newbury Park, CA, 1985.

Patton, M.Q. *Qualitative Evaluation and Research Methods* Sage Publications, Newbury Park, CA, 2002.

Poe, V., Brobst, S., and Klauer, P. *Building a Data Warehouse for Decision Support* Prentice Hall, NJ, USA, 1997.

Ponniah, P. *Data Warehousing Fundamentals: A Comprehensive Guide for IT Professionals* Wiley, New York, 2001.

Sen, A. "Metadata Management: Past, Present and Future," *Decision Support Systems* (37:1), 2002, pp 151-173.

Shankaranarayanan, G. "Towards Implementing Total Data Quality Management In A Data Warehouse," *Journal of Information Technology Management* (16:1) 2005, pp 21-30.

Shankaranarayanan, G., and Even, A. "The Metadata Enigma," *Communications of the ACM* (49:2) 2006, pp 88-94.

Shankaranarayanan, G., Wang, R.Y., and Ziad, M. "IP-MAP: Representing the Manufacture of an Information Product," in: *Proceedings of the 2000 International Conference on Information Quality*, 2000.

Staudt, M., Vaduva, A., and Vetterli, T. "Metadata Management and Data Warehousing." Technical Report 21, Information Systems Research, July 1999, University of Zurich.

Tan, W., Shen, W., Xu, L., Zhou, B., and Li, L. "A Business Process Intelligence System for Enterprise Process Performance Management," *IEEE Transactions On Systems, Man, And Cybernetics - Part C: Applications And Reviews* (38:6) 2008.

Vaduva, A., and Dittrich, K.R. "Metadata Management for Data Warehousing: Between Vision and Reality," in: *2001 International Symposium on Database Engineering & Applications*, Grenoble , France, 2002.

Viaene, S. "Linking Business Intelligence into Your Business," *IEEE Xplore Computer Journal* (10:6) 2008, pp 28-34.

Wang, R. "A Product Perspective on Total Data Quality Management," *Communications of the ACM* (41:2) February 1998, pp 58-63.

Wang, R., Lee, Y., Pipino, L., and Strong, D. "Manage Your Information as a Product," *Sloan Management Review* (39:4) Summer 1998, pp 95-105.

Wang, W., Ding, H., Dong, J., and Ren, C. "A Comparison of Business Process Modelling Methods," in: *IEEE International Conference on Service Operations and Logistics, and Informatics Proceedings*, Shanghai, China, 2006.

Weske, M. *Business Process Management: Concepts, Languages, Architectures* Springer, Berlin, 2007.

White, S.A. "Introduction to BPMN," IBM Corporation, 2004.

Wixom, B., and Watson, H. "The BI-Based Organization," *International Journal of Business Intelligence Research* (1:1) 2010, pp 13-28.

Yeoh, W., and Koronios, A. "Critical Success Factors for Business Intelligence Systems", *Journal of Computer Information Systems*, (50:3), 2010, pp 23-32.

Yin, R.K. *Case Study Research: Design and Methods*, (3rd ed.) Sage Publications, Thousand Oaks, CA, 2003.

APPENDIX

Table 2. Third Layer of BIP-Map for Job Advertisement Graph (ETL Stage)

Database		ETL Process	Data Warehouse	
advertisement Table		1. Create the **DimIndustry** dimension table. • DimIndustry.IndustryKey: Auto-generated • ref_industry.industry_id → DimIndustry.IndustryID • ref_industry.industry_field → DimIndustry.IndustryField 2. Create the **DimAdvertiser** dimension table. • DimAdvertiser.AdvertiserKey: Auto-generated • advertiser.advertiser_code → DimAdvertiser.AdvertiserCode 3. Create the **DimPostingDate** dimension table and **FactJobAdvertisement** fact table. • DimPostingDate.PostingDateKey: Auto-generated • advertisement.posting_date → DimPostingDate.PostingDate • YEAR(advertisement.posting_date) → DimPostingDate.PostingYear • DATENAME(MONTH, advertisement.posting_date) → DimPostingDate.PostingMonth • DAY(advertisement.posting_date) → DimPostingDate.PostingDay • DimAdvertiser.AdvertiserKey → FactJobAdvertisement. AdvertiserKey • DimIndustry.IndustryKey → FactJobAdvertisement.IndustryKey • DimPostingDate.PostingDateKey → FactJobAdvertisement. PostingDateKey • advertisement.advertisement_code → FactJobAdvertisement.AdvertisementCode • advertisement.advertisement_status → FactJobAdvertisement. AdvertisementStatus	**FactJobAdvertisement Table**	
advertisement_code	int		AdvertiserKey	int
advertiser_code			IndustryKey	
posting_date	datetime		PostingDateKey	
advertisement_status	int		AdvertisementCode	
			AdvertisementStatus	
advertiser Table				
advertiser_code	int		**DimAdvertiser Table**	
industry_id			AdvertiserKey	int
			AdvertiserCode	
ref_industry Table				
industry_id	int		**DimIndustry Table**	
industry_field	varchar(255)		IndustryKey	int
			IndustryID	
			IndustryField	varchar(255)
			DimPostingDate Table	
			PostingDateKey	int
			PostingDate	datetime
			PostingYear	int
			PostingMonth	varchar(50)
			PostingDay	int

Table 3. Third Layer of BIP-Map for Job Advertisement Graph (Data Warehouse)

FactJobAdvertisement					
Data	**Description**	**Input Type**	**Data Source**	**Data Type**	**Data Code**
AdvertiserKey	– A number used to uniquely identify an advertiser.	Extraction	– DimAdvertiser * AdvertiserKey	int	-
IndustryKey	– A number used to uniquely identify an industry.		– DimIndustry * IndustryKey		-
PostingDateKey	– A number used to uniquely identify a job posting date.		– DimPostingDate * PostingDateKey		-
AdvertisementCode	– A number used to uniquely identify a job advertisement.		– advertisement * advertisement_code		-
AdvertisementStatus	– Indicates whether the job advertisement is posted online or not. – Job advertisement that is saved but not posted online will not be displayed.		– advertisement * advertisement_status		1 : Posted Jobs 2 : Not Posted Jobs

DimIndustry					
Data	**Description**	**Input Type**	**Data Source**	**Data Type**	**Data Code**
IndustryKey	– An auto-generated number used to uniquely identify an industry.	Generation	– Auto-Generated	int	-
IndustryID	– The original number used to uniquely identify an industry.	Extraction	– ref_industry * industry_id		-
IndustryField	– A name that describes the industry.		– ref_industry * industry_field	varchar(255)	-

DimAdvertiser					
Data	**Description**	**Input Type**	**Data Source**	**Data Type**	**Data Code**
AdvertiserKey	– An auto-generated number used to uniquely identify an advertiser.	Generation	– Auto-Generated	int	-
AdvertiserCode	– The original number used to uniquely identify an advertiser.	Extraction	– advertisement * advertisement_code		-

DimPostingDate					
Data	**Description**	**Input Type**	**Data Source**	**Data Type**	**Data Code**
PostingDateKey	– An auto-generated number used to uniquely identify a job posting date.	Generation	– Auto-Generated	int	-
PostingYear	– The year of a job posting date.	Transformation	– advertisement * posting_date		-
PostingMonth	– The month of a job posting date.			varchar(50)	-
PostingDate	– The day of a job posting date.			int	-