DATA DEPENDENCY TO THE ORGANIZATIONAL GOALS: A CONCEPT AND FRAMEWORK

Trieu Minh Nhut Le
Department of Computer Science and Computer Engineering
Saigon University
Ho Chi Minh City, Vietnam

Received date: 3 April 2018
Accepted date: 2 August 2018

ABSTRACT

In the context of the organizational goals, evaluation of data linkage is essential to identify dependencies of organizational data to the organizational goals that may reflect on decision making process. The issue is to identify relevance of organizational data from a large number of organization data sets. This paper attempts to identify the ordering of achievement of organizational goals based on analysis of dependencies of organizational data. In this paper we extend the development of the proposed organizational goals ontology by looking at the evaluation approach which can incorporate the dependencies of organizational data to the organizational goals. The aim of this paper is to evaluate the weight of this dependency at design level and operational level.

.

Keywords: data dependency, data goals dependency, evaluation levels, frequency value, organizational goals

INTRODUCTION

Government agencies, public and private bodies are drowning in an ever-increasing deluge of data (Christen, 2012) because they create and collect massive amount of data in their daily business activity. Thus, the ability to analyse their data in a timely fashion can provide a competitive

edge to improve productivity in relation to the organizational goals. Data is the most important asset to assist the decision making process in achieving the organizational goals. However, the trustworthiness of organizational data in relation to the organizational goals is always questionable due to the huge data mining issue within the organization. Some of this data become are not relevant or over time become irrelevant to the organizational goals. Therefore, it is difficult to identify the relevance of organizational data even though the professional such as data analysts are trained to analyse this data but the increase amount of organizational data has become a major problem in applying this data for achieving the organizational goals. The growth in the amount of the organization resources available nowadays poses major difficulties as well as challenges in decision making (Mikroyannidis & Theodoulidis, 2010). Thus, modelling the organizational goals structure is important to develop a relationship between organizational data and organizational goals. For example, an ontology improves the understanding of the organizational goals structure as it shows the relationship of the organizational goal elements (Izhar, Torabi, Bhatti, & Liu, 2012; Izhar, Torabi, Bhatti, & Liu, 2013).

In data retrieval process, data is classified in different data sets. For example, data sets for each organization department or data sets for staffs and customers. These data sets stored a vast amount of data and some of this data become not relevant. As a result, it creates a challenge to identify relevance of data to support dynamic decision making in relation to the organizational goals. Given this issue, the past decade has seen strong interest in new approach that allows the efficient data processing, analysis and management of large data collection such as data warehouse and data mining. Data warehouse is process used for reporting and data analysis as a central repository of data to integrate data from various data sources (Romero & Abello, 2010; Selma et al., 2012). Data mining is a process of discovering new information from a vast data collection (Mansingh, Osei-Bryson, & Reichgelt, 2009). To our knowledge, there is no study consider the evaluation of organizational data to assist decision making process in relation to the organizational goals. Modelling the organizational goals limited to the business process and the organization process (Fox, Barbuceanu, Gruninger, & Lin, 1998; Mansingh et al., 2009; Rao, Mansingh, & Osei-Bryson, 2012; Sharma & Osei-Bryson, 2008). Therefore, it is important to identify the dependency between organizational data and organizational goals.

Despite of this shortfall, there is shortcoming when it comes to

data linkage (Durham, Xue, Kantarcioglu, & Malin, 2012; Freire et al., 2012; Meray, Reitsma, Ravelli, & Bonsel, 2007). However, past studies are limited to data linkage software (Ferrante & Boyd, 2012; Trepetin, 2008) and database (Freire et al., 2012; Su, Wang, & Lochovsky, 2010). In organization, structuring a small organization is less complicated compare to a large organization. Different organization structures, processes, amount of data attributes make it more difficult to identify the relevance of organizational data in relation to the organizational goals. Therefore, it is important to identify the dependency of organizational data that relate to the organizational goals. In this paper, our main contribution is to extend the concept of organizational goals (Izhar et al., 2012; Izhar et al., 2013) by looking at the evaluation approach which can incorporate the dependency between organizational data that relate to the organizational goals. In addition, we suggest that organizational data that have the higher value of number to the organizational goals is considered relevant in relation to the organizational goals.

In contrast to our studies in Izhar et al. (2012) and Izhar et al. (2013), we tackle the problem of evaluating organizational data that relate to the organizational goals based on the following aims. The aim of this paper is to evaluate the dependencies of organizational data that relate to the organizational goals ontology. The evaluation is based on two levels; design level and operational level. Firstly, design level is to build a conceptual map of data elements and we identify the dependency of this data in relation to the organizational goals. We suggest design level is the evaluation that based on the organizational data sets. The process is important to identify data attributes that relate to the organizational goals. Secondly, operational level is to evaluate the respondent of the actual data instance and frequency of data are analysed. In this paper, the methodology is proposed to comparatively evaluate the value of organizational data at design level and operational level. We suggest that it is important to measure the value of data attributes as an effort to consider this data is relevant in achieving the organizational goals. In order to identify the strength measurement for the relevance of organizational data, one approach is to analyze the frequency value of organizational data in relation to the organizational goals. At this stage, we suggest that frequency value as a metrics to measure the weight of organizational data that relate to the organizational goals. In this paper, data goals dependency is a task to identify the dependency relationship of organizational data from vast data sets as the contribution of this paper

will serve as a first step in evaluating organizational data that relate to the organizational goals.

Motivation

This paper is present to explain in detail, along with how the proposed organizational goals ontology is implemented and applied to evaluate organizational data that relates to the organizational goals (Izhar et al., 2012; Izhar et al., 2013). This paper is present to validate the flexibility of the proposed model to assist domain experts and entrepreneurs for decision making process in relation to the organizational goals. Meanwhile, the experiment evaluation is presented to discover any issues or gaps which may be presented during the implementation at design and operational levels. This is considered as flexibility to confirm the validity of the proposed model that can be applied in any situation. This paper specify the organizational resource which is organizational data as it is suggested that organizational data is the most important organizational resource in relation to the organizational goals (Izhar et al., 2012; Izhar et al., 2013). The outcome of this paper will improve the process of evaluating organizational data that relates to the organizational goals.

To achieve this, the main challenge is to identify organizational data from organizational data sets that relate to the organizational goals. This paper is present to identify the dependencies of organizational data in relation to the organizational goals which is design level and to evaluate the weight of organizational data that relate to the organizational goals which is operational level. The incentive outcome of this paper covers the limitations as follows:

1. to successfully identify organizational data that relate to the organizational goals,

2. to successfully evaluate the weight of organizational data that relate to the organizational goals,

3. to proposed model that is flexible and scalable so that it can be applied with any organizational data in relation to the organizational goals.

The outcome of this paper will improve the process of evaluating organizational data that relate to the organizational goals. This paper is aim to advance the understanding of the dependency relationship of organizational data and organizational goal. At the mean time, the contribution of this paper will serve as an introduction to evaluate the weight of organizational data that relate to the organizational goals as this organizational data is consider relevant in achieving the organizational goals.

RELATED WORK

Recently, there have been substantial growths in record linkage activities (Durham et al., 2012; Freire et al., 2012; Meray et al., 2007). Most of these studies focused on the task of identifying data from data sets in order to prevent any redundancy of data. To our knowledge, there is no study has been carried out in the development of organizational data in relation to the organizational goals. Even though study on organizational goals has been carried out but most of the studies focus of the modelling concept for organization performance (Bouskila-Yam & Kluger, 2011; Earley & Kanfer, 1985; Lepmets, McBride, & Ras, 2012; Salerno, 2009). Therefore, it is important to identify the link between organizational data and organizational goals as we suggest this data should be relevant to assist any decision making process for the achievement of the organizational goals.

Record linkage

Record linkage defined as a task performed to identify record from number of data sets (Christen, 2008; Durham et al., 2012). Past studies on record linkage are mostly focus on medical records (Freire et al., 2012; Jutte, Roos, & Brownell, 2011; Meray et al., 2007) or software (Ferrante & Boyd, 2012; Trepetin, 2008). Meray et al. (Meray et al., 2007) focused on medical record. They applied probabilistic record linkage technique to combine databases without a patient identification number. This technique allowed the creation of a high-quality linked database from data sets. In this study, the authors suggest the technique is useful for linkage of any anonymous registries in the absence of personal identifiers and goal standard. Another example of record linkage in medical record is Freire et al. (Freire

et al., 2012). The authors look at record linkage process in screening patient database because the patient information is not uniquely identified. For example, record linkage is presented to integrate the file in database called Brazillian Cervical Cancer Information System (SISCOLO). The authors show that record linkage integrate SISCOLO to produce indicators for the evaluation of the cervical cancer screening programme taking the patient as a unit of observation. Record linkage assesses the effectiveness and coverage the quality of data as a way to contribute to a more efficient use of SISCOLO in the planning of health actions.

At the same time, data privacy also becomes an issue when implementing record linkage (Karakasidis & Verykios, 2011). The authors propose a technique to address the problem of efficient privacy preserving approximate record linkage. The technique is important to combine the speed of private blocking with the increased accuracy of approximate secure matching of data. In this study, the technique did not apply any data that match to any organizational goals but it is more on the speed to retrieving the data. Abril et al. (2012) also studied on record linkage in the context of data privacy. The authors suggest that record linkage can be used as an estimator of the disclosure risk of protected data. During the linkage process, the authors introduce a parameterization of record linkage to improve the linkage as compared to standard distance-based record linkage and to identify the key-attributes for record linkage. Therefore, the authors determine the weight identification for every linkage process, which express the importance of each variable in the linkage process. Thus, in data privacy, record linkage is used as a disclosure risk estimation of the protected data.

Christen (2012) looked at the indexing techniques for scalable record linkage and de-duplication of data. The author argued that matched data are becoming important in many application areas, because they can contain information that is not available otherwise, or that is too costly to acquire. Removing duplicate records in a single database is a crucial step in the data cleaning process, because duplicates can severely influence the outcomes of any subsequent data processing or data mining. In contrast, this study is important as we attempt to prevent any redundant data for data analysis in relation to the organizational goals. Previously, Christen (2008) studied on an automatic record linkage using seeded nearest neighbour and support vector machine classification that aimed at automating the record linkage process.

Ferrante & Boyd (2012) propose a transparent and transportable

methodology for evaluation of record linkage software. The methodology use to evaluate record linkage software to improve the quality linkage. The authors evaluate large number of package that involves the use of synthetic data using pre-defined linkage strategy and the use of standard linkage quality metrics to assess the performance. They suggest that methodology as an unique opportunity to benchmark the quality of linkage in different operational environments. Another example is Durham et al. (2012). The authors study on quantifying the correctness, computational complexity and security of privacy-preserving string comparators for record linkage. In this study, privacy-preserving record linkage is a variant of the task in which data owners wish to perform linkage without revealing identifiers associated with the records. It shows the privacy of data that can keep the security of the customer data privacy. Even though studies on data linkage have been carried out, to our knowledge, there is no study link the relationship between organizational data that relate to organizational goals. Previous approach on metrics evaluates the weight of the linkage process (Varghese & Sundar, 2011). We adapted metrics evaluation that used with various types of databases (Herzog, Sheuren, & Winkler, 2007). In contrast, one approach to measure the weight is to identify the frequency value as a metrics to identify relevance of each data attribute that relate to the organizational goals.

Organizational goals

Organizational goals are defined as the organization main target. It is the higher and important achievement target in every organization and it consist the process of identifying the aim of the organization. Thus, it is important to understand the organizational goals structure. The structure of the organization is important to develop the efficiency and flexibility of the organization to cope with unpredictable (Salerno, 2009). For example, organization structure is developed to achieve the performance of the organization (Bouskila-Yam & Kluger, 2011) or the goal structure is developed to achieve the performance of goals (Barlas & Yasarcan, 2006; Earley & Kanfer, 1985; Lepmets et al., 2012; Sholihin, Pike, Mangena, & Li, 2011). The example shows a number of the studies that look at the organization structure toward the performance. This is because the organization performance depends on the organization structure. Same with

the goal structure and the goal performance, in which the organizational goals depend on the goal structure toward the goal performance. There are number of the organizational goals studies that focus on the performance such as system performance (Ceresia, 2011; Kang & Norton, 2004), goal performance (Barlas & Yasarcan, 2006; Dillard, 1981; Sholihin et al., 2011) and organization performance (Bouskila-Yam & Kluger, 2011; Earley & Kanfer, 1985; Lepmets et al., 2012; Salerno, 2009). It is important to identify the entire organization modelling process as an effort to look at the organization performance and the goal performance. However, the process can be very large and it is very difficult to evaluate the organizational data as an effort to achieve the organizational goals. In contrast, it is important to identify organizational data that relate to the organizational goals. Therefore, the linkage of organizational data from data sets should be consistent in order to prevent any redundant data, thus it can be evaluated in relation to the organizational goals.

Even though the concept of the organizational goals is developed but modelling the structural of the organizational goals always questionable. Thus, we suggest that ontology is important to develop a common understanding of the organizational goals structure (Izhar et al., 2012). At the same time, ontology is explicit and formal specifications of the knowledge, especially implicit or hidden knowledge (Cho, Han, & Kim, 2006). Ontology also considered as an approach to support data sharing (Pundt & Bishr, 2002). Thus, ontology assists with part of the integration problem in relation to the organizational goals. Therefore, ontology can be used to improve communication between decision makers and users collaborating (Selma et al., 2012), where in our case, the communication between the decision makers in relation to the organizational goals.

The concept of data dependency to the organizational goals

The aim of this paper is to implement the proposed model in Izhar et al. (2012) and Izhar et al. (2013) to evaluate organizational data that relate to the organizational goals as an effort to consider this organizational data as relevant data in relation to the organizational goals. In order to achieve this aim, we proposed an ontology as an approach to identify the dependency relationship of organizational goals element which includes sub-goal and organizational data (Izhar et al., 2012; Izhar et al., 2013). We suggest an

ontology is important to improve the understanding of the dependency relationships of the organizational goals elements. Despite the vast amount of research on ontology, there remains relatively little research in applying an ontology for the organizational goals. In our work (Izhar et al., 2013), we unified previous models that cater the structure of the organizational goals (Fox et al., 1998; Rao et al., 2012; Sharma & Osei-Bryson, 2008). Fox et al. (1998) addressed an organization to be a set of constraint on the activities performed by the organizations. This study has been extended in Sharma & Osei-Bryson (2008), who discussed the relationships such as the business process with a various types of the resources. Sharma & Osei-Bryson (2008) and Rao et al. (2012) included the entire organization resources in their ontology.

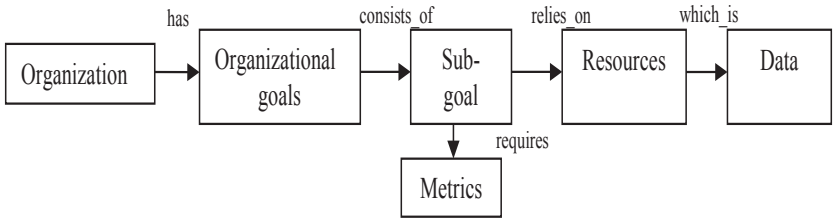

Fig. 1. Organizational goals ontology.

In contrast to the previous studies (Fox et al., 1998; Rao et al., 2012; Sharma & Osei-Bryson, 2008), we evaluate the organization resource which is organizational data. In this paper, we do not interpret the overall business process that includes the overall organization activities and performance. Instead of evaluating the information and knowledge, we argue that the organizational data is important and must be considered during the decision making process in relation to the organizational goals (Izhar et al., 2012; Izhar et al., 2013) because it is very important to identify the relevant of organizational data in order to achieve the organizational goals. Most of the previous studies focused on the organization structure and performance which include the entire business processes and organization processes (Fox et al., 1998; Rao et al., 2012; Sharma & Osei-Bryson, 2008). We adapted these studies in the organizational goals ontology as shown in Fig. 1.

Fig. 1 is the extended version of the organizational goals ontology (Izhar et al., 2012; Izhar et al., 2013). Fig. 1 shows the relationships of organizational goal elements. However, we attempt to evaluate organization resource which is organizational data in relation to the organizational goals. We suggest it is important to identify organizational data that relate to the organizational goals. Therefore, we can consider this data is relevant in relation to the organizational goals. It provides domain experts and entrepreneurs with knowledge to identify the most relevant organizational data in relation to organizational goals. One approach to evaluate this data is using a metrics as a measurement approach to evaluate the weight of organizational data that relate to the organizational goals.

Data goals dependency

Even though there are many studies have been carried out in the context of the data process (Kum, Duncan, & Stewart, 2009; Liao, Chang, & Lee, 2008), limited study has been observed in evaluating organizational data in relation to the organizational goals (Izhar et al., 2012; Izhar et al., 2013). Therefore, it is important to identify the dependency relationship between organizational data and organizational goals. The relationship is important to identify the relevance of organizational data from data sets in achieving the organizational goals. However, organizations have a huge set of the organizational data that might be relevant to the organizational goals. This large set of the organizational data might not be relevant with respect to the organizational goals. Thus, the first step to identify the relevance of organizational data is to recognize the dependency of organizational data for the achievement of the organizational goals.

To achieve this aim, record linkage approach is adapted to identify the possible relationship between organizational data and organizational goals. Record linkage is highly used to identify data that being linked, so all data sets under consideration should ideally undergo a matching process prior to the record linkage (Durham et al., 2012). Even though there are studies have been carried out in various issues such as software (Freire et al., 2012; Jutte et al., 2011), data privacy (Abril et al., 2012; Karakasidis & Verykios, 2011) and security (Durham et al., 2012), we suggest it is

important to develop a standard set of approach to show the relationship between data attributes and organizational goals. Even though the term is defined as record linkage, data linkage, record matching, data matching (Abril et al., 2012; Christen, 2012; Ferrante & Boyd, 2012; Scannapieco, Figotin, Bertino, & Elmagarmid, 2007; Su et al., 2010; Yakout, Elmagarmid, Elmeleegy, & Ouzzanil, 2010), but we use the term data dependency as an effort to identify the dependency relationship between data attributes and organizational goals because we attempt to identify the dependency for each data attributes that relate to the organizational goals. We propose data goals dependency based on the organizational goals ontology (Izhar et al., 2012; Izhar et al., 2013) as we define data goals dependency is a process to identify the existing organizational data attributes from the organizational data sets in relation to the organizational goals as shown in Fig. 2.

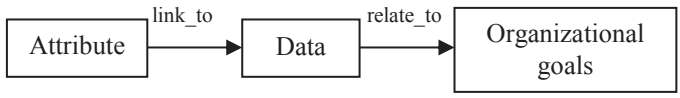

Fig. 2. Data goals dependency.

METHODOLOGY

The scope of the methodology consists on the development approach for domain experts and entrepreneurs to access available organizational data for the decision making process. As part of the organizational goals ontology development, we introduce an approach steps to specify the process of identify relevance of organizational data from data sets that relate to the organizational goals. In contrast to the works in Rao et al. (2012), Sharma & Osei-Bryson (2008) and Fox et al. (1998), methodology in this paper will propose an evaluation steps to identify the dependency relationship between organizational data and organizational goals and to evaluate the weight of this dependency so this organizational data can be considered relevant. The methodology is important to specify to what extend the organizational goals are achieved. Therefore, future approach can be suggested to improve any gaps and issues in achieving the organizational goals.

The development of the methodology consists of two stages. Firstly to specify dependency relationship of the organizational goals and how this organizational data relate to the organizational goals as discussed in Section 3. Secondly to develop a measurement approach as a metrics to analyse organizational data so this data can be considered relevant in achieving the organizational goals. The aim of this methodology is to measure organizational data that relates to the organizational goals as an effort to assist domain experts and entrepreneurs for decision making process in achieving the organizational goals. This methodology is design into three main approaches as shown in Fig. 3.

1. Firstly, identify dependency relationship of the organizational goals based on ontology (Izhar et al., 2013). Ontology is important to define a specification of a conceptualization. Ontology is developed to categories some domain within the concept and it is important to improve the understanding of the structure of data or information among people or software.

2. Secondly, identify dependency relationship of organizational data that relates to the organizational goals. It is important to identify this relationship of organizational data that relates to the organization goals as the first step to identify data attributes to be analysed in relation to the organizational goals.

3. Thirdly, the development of a metrics approach as an analysis tool to measure organizational data to be considered relevant in relation to the organizational goals.

The trustworthiness of organizational data that relates to the organizational goals is important as this organizational data is consider relevant to assist domain experts and entrepreneurs for decision making process in relation to the organizational goals. In Fig. 3, assuming the process of organizational goals involves three sub-goals but in real case organization might have more than three sub-goals. In this figure, it is important to identify the value for each data attribute that relate to the different sub-goals so this data attributes can be considered relevant to the organizational goals.

Therefore, metrics definition is important for this measurement analysis so the value can be analysed based on the requirement of domain experts and entrepreneurs for decision making process in relation to the organizational goals. The process of measurement analysis become a nature of decision making as the analysed data present the value of organizational data that relate to the organizational goals.

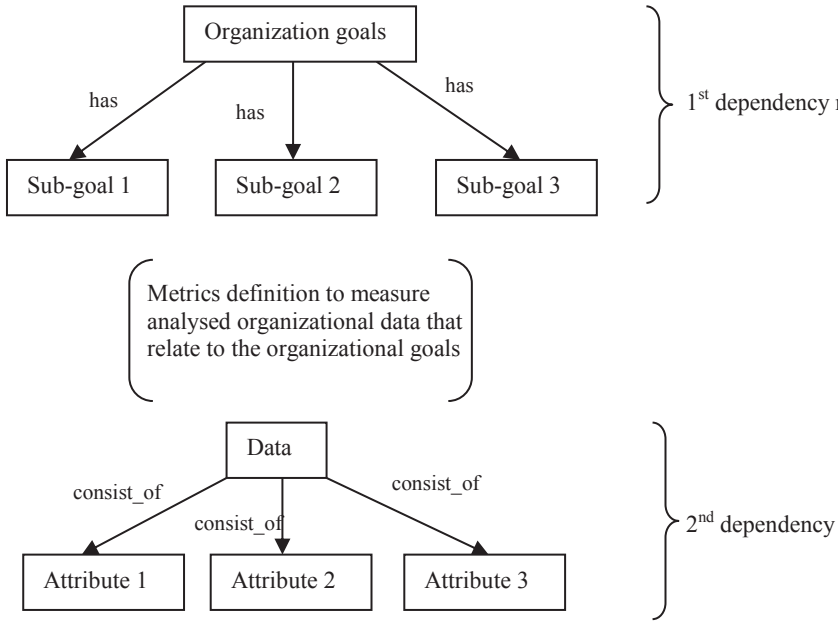

Fig. 3. Concept for the methodology.

The first half of this paper is to develop the dependency of organizational data from data set that relate to the organizational goals based on the proposed data goals dependency as we suggest this process is design level. Design level is proposed to build a conceptual map of data elements and we identify dependency relationship of this data in relation to the organizational goals. At design level, we attempt to identify the possible organizational data attributes that relate to the organizational goals.

As an effort to evaluate organizational data that relate to the organizational goals, it is important to measure the value to this data so it can be considered relevant in relation to the organizational goals. Therefore, we define this process as an operational level. Operational level is defined as a process to evaluate the respondent of the actual data instance and frequency of data are analysed. At operational level, measurement metrics is important to identify the value for analysed data in relation to the organizational goals. In order to identify the strength measurement for the relevance of organizational data, one approach is to analyze the frequency value for each data attributes. At this stage of our research, we consider frequency value as a metrics to measure a weight dependency for organizational data that relate to the organizational goals as shown in Fig. 4.

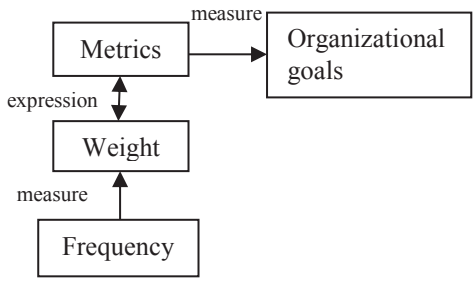

Fig. 4. Goals metrics for the measurement of data goals dependency at operational level.

Metrics definition to the organizational goals

In this section, we attempt to measure organizational data that relates to the organizational goals. At this stage we identify metrics approach to measure analysed data at the operational level for domain experts to come out with decision making process in relation to the organizational goals. Fig. 2 showed the dependency relationship of organizational data and organizational goals. It is important to identify the value for each data attributes that relates to the organizational goals so this organizational data can be considered as relevance of organizational data in relation to

the organizational goals. Therefore, it is important to define the metrics assuming the factor of;

- weight of the dependency relationship,
- number of data attributes,
- frequency of data attributes,
- percentage value of data attributes.

In general, when the value is applied, the value of weight will change and this value can be defined in many ways such as percentage or frequency value based on different situation. For example, domain experts and entrepreneurs might want to identify the percentage value of selected data or they might want to identify the average value of selected that relates to the organizational goals. Assuming this value is analysed and it can be presented on the dashboard to show a graphical presentation of value, the comparison of this value can be presented to support decision making process in relation to the organizational goals.

Previous approach in metrics evaluates the weight of the linkage process of matching data (Varghese & Sundar, 2011) for metrics evaluation that used with various types of databases (Herzog et al., 2007). In this paper, we attempt to evaluate the dependency relationship of organizational data in relation to the organizational goals.

Frequency value

We define frequency is the number of times a value from gathered data is relate to the organizational goals. At this stage, we consider frequency value as a metrics to measure the weight of the relevance of data attributes in which we defined as number of data that relate to the organizational goals. In contrast to Ferrante et al. (Ferrante & Boyd, 2012), the authors did not defined weight that be used in linkage of data. However, we suggest that it is important to specify weight in the dependency relationship between organizational data and organizational goals because it will assists decision

maker to make decision in relation to the organizational goals based on the relevance of organizational data. At the same time, we suggest datasets can be organized to show the frequency of organizational data that occur in achieving the organizational goals. Therefore, we attempt to present this value as an effort to identify relevance of organizational data that relate to the organizational goals at the operational level.

In this paper, we attempt to identify which data attributes relate to the organizational goals and to what extent this organizational data is relevance to the organizational goals. In order to achieve this aim, one approach is to identify the frequency value of this data. In Section 3, we proposed data goals dependency to identify organizational data in relation to the organizational goals. At this stage, we attempt to analyze the weight of frequency for organizational data that relate to the organizational goals as we define the weight of this frequency as:

Frequency value=((Number of value)/(Total number of value))

Let's come out with an example to support our explanation as shown in Table 1. In this example, assuming we identified university scholarship among student that have not used the services. This table show number of university scholarship that relates to postgraduate as goal 1 and undergraduate as goal 2. In this example, if number of value for university scholarship is 490 for goal 1 and 2357 relate for goal 2. The total number of value is 2847. Therefore, we conclude the frequency value of university scholarship to goal 1 is 490/2847 and frequency value of university scholarship to goal 2 is 2357/2847.

Table 1 Example of frequency value for university scholarship.

Study level	Postgraduate		Undergraduate	
	Number	Frequency value	Number	Frequency value
University scholarship	490	0.172111	2357	0.827889

In Christen (2008), the author discussed matching weight between 1 and 0 and all matching weight calculated of a pair of data will be 1. It is a typically captured via a similarity function that given a pair of strings returns a number of between 0 and 1 (Arasu, Chaudhuri, & Kaushik, 2008). At this stage, we computed the frequency of total data that relate to the organizational goals. Therefore, the closer value to 1, the more relevant this data and this value should complementary with the value 1. We define the value as:

min(value), close to 0

max(value), close to 1

as, $1 \geq a \geq 0$.

EVALUATION LEVELS

The evaluation in this paper can be concluded based on design level and operational level. Design level is a conceptual map of data elements and we identified the dependencies of this data in relation to the organizational goals as the methodology in Section 4 proposed an approach for domain experts and entrepreneurs to follow to assist their decision making process in relation to the organizational goals. In design level, we identified the possible variables that relate to the organizational goals. It is important to apply the methodology at the operational aspect. Meaning, we need to identify the weight of analysed data. It is based on the actual implementation in which from the student feedback and the actual respond from the experiment. The evaluation results only show the summary of data dependency of organizational data and organizational goals. Based on the discussion, the process is summarized for the evaluation stage of the methodology. These levels (design and operational) are not part of the methodology as the discussion only to advance the understanding on how the methodology could be evaluated and implemented.

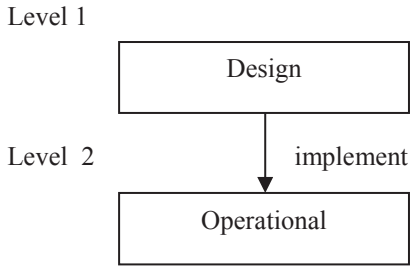

Fig. 5. Evaluation levels.

Design level

In design level, the conceptual model as defined in the methodology is a conceptual phase to define the organizational goals. It is also define as a conceptual scope for the organizational goals. It is a process of defining the dependency relationship of the organizational goals. The process looks at the possible number of goals in organization and how organizational data relate to these goals. In Section 4, the process includes the dependency relationship of the organizational goals and dependency relationship between organizational data and organizational goals.

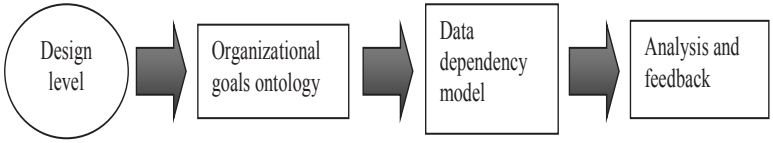

Fig. 6. Design level.

In Fig. 6, we see the methodology involve the step of organizational goals ontology and data dependency which mean data dependency between organizational data and organizational goals. Design level summarized entire idea how the methodology is developed based on organizational goals scope. At the same time, it is the process of defining the goals in real practical case as discussed in the next section.

Operational level

In operational level, it is a process of evaluating presented data in the experiment in which relate to the experiment goals. It is a set of metrics based on the design level (conceptual scope for the organizational goals) which associated with the organizational goals in order to achieve it in a measurement way. It is a process of identifying the weight of organizational data that relate to the organizational goals.

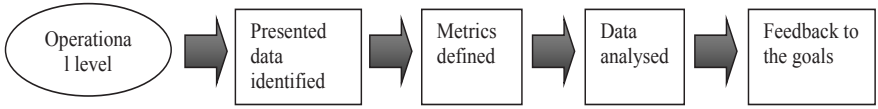

Fig. 7. Operational level.

In Fig. 7, we see the process on data analysis that relate to the organizational goals. In is the process after experiment goals have been identified, the dependency relationship developed and the process show how the presented data is analyse. It is the level that discussed how data is evaluated and how this evaluated data and it value is presented to the goals. In decision making, the analysed value could be presented in dashboard as one example for domain experts and entrepreneurs to compare the most achievable value toward the goals.

DISCUSSION

The dependency process presented in this paper was attempted to identify relevance of organizational data from organizational datasets in relation to the organizational goals. Although study on data linkage has been carried out (Durham et al., 2012; Freire et al., 2012; Meray et al., 2007), to the best of our knowledge, there is no study discuss on data linkage in the context of the organizational goals. In this paper, we evaluated organizational data that relate to the organizational goals based on two levels; design level and operational level.

At design level, it is important to identify data and attributes from data sets (e.g. data collection from survey). The process is important to identify data that relate to the organizational goals. At operational level,

the evaluation is based on the actual implementation of data collection. In this level, we attempted to identify the frequency value as a metrics of data attributes in relation to the goal 1 and goal 2.

We suggest data goals dependency improve the decision making process in relation to the organizational goals. Data goals dependency shows the dependency relationship between organizational data and organizational goals. At the same time, data goals dependency assists the process of identifying data attributes, where we suggest that these data attributes are relevance in relation to the organizational goals. However, there is still limitation because we applied small structure of organization data collection. In this paper we used data from La Trobe University Student Support Service Survey and we evaluate this data in relation to the survey's goal instead of the entire university's goals. Data structure, volume of the data and business process are difference for the organizations. At the same time, we will extend the approach for data evaluation because in the future, other approach such as ranking system can be used to assist the process of identifying the frequency value. In this paper, we suggest frequency value as a metrics to measure the weight for data dependency that relate to the organizational goals. We suggest measuring the frequency value is important to sort the value of data analysed as an effort to improve decision making process in relation to the organizational goals.

CONCLUSION

In this paper we have presented and analysed data goals dependency based on data linkage. This is done by identifying organizational data that relate to the organizational goals. We have provided an approach of identifying the frequency value of data that relate to the organizational goals.

In the first half of this paper, we discussed the concept of the data linkage in the scope of the organizational goals. In order to develop the data goals dependency, we generalize organizational data from data sets in relation to the organizational goals. In this paper, we applied data from La Trobe University Student Support Service survey. We presented data dependency from the survey in relation to the La Trobe University Student Support Service survey's goals. In the second half of this paper, we evaluated data goals dependency at design level and operation level. The evaluation estimates the frequency value as a metrics for data attributes that relate

to the organizational goals. We conclude that frequency value for data is important to identify which organizational goals are more important and achievable. Our results show data goals dependency assist the process of evaluating data for decision maker to make any decision making based on analysed data in relation to the organizational goals.

In the conclusion, it is important to identify the dependency relationship between organizational data and organizational goals for the achievement of the organizational goals. Data is the most important organization resource and the organization can embrace the fact that it is important to have a relevance of data to achieve the organizational goals or facing the problem in identifying the relevance of data from the vast amount of it.

REFERENCES

Abril, D., Navarro-Arribas, G., & Torra, V. (2012). Improving record linkage with supervised learning for disclosure risk assessment. *Information Fusion, 13*(4), 274-284. doi: http://dx.doi.org/10.1108/02635571111133542

Arasu, A., Chaudhuri, S., & Kaushik, R. (2008). *Transformation-based framework for record matching*. Paper presented at the IEEE 24th International Conference, Cancun, Mexico.

Barlas, Y., & Yasarcan, H. (2006). Goal setting, evaluation, learning and revision: A dynamic modeling approach. *Evaluation and Program Planning, 29*(1), 79-87.

Bouskila-Yam, O., & Kluger, A. N. (2011). Strength-based performance appraisal and goal setting. *Human Resource Management Review, 21*(2), 137-147.

Ceresia, F. (2011). A model of goal dynamic in technology-based organizations. *Journal of Engineering and Technology Management, 28*(1-2), 49-76.

Cho, J., Han, S., & Kim, H. (2006). Meta-ontology for automated information integration of parts libraries. *Computer-Aided Design, 38*(7), 713-725.

Christen, P. (2008). *Automatic record linkage using seeded neareast neighbour and support vector machine classification.* Paper presented at the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Las Vegas, Nevada, USA.

Christen, P. (2012). A survey of indexing techniques for scalable record linkage and deduplication. *IEEE Transaction on Knowledge and Data Engineering, 24*(9), 1537-1555.

Dillard, J. F. (1981). A longitudinal evaluation of an occupational goal-expectancy model in professional accounting organizations. *Accounting, Organizations and Society, 6*(1), 17-26.

Durham, E., Xue, Y., Kantarcioglu, M., & Malin, B. (2012). Quantifying the correctness, computational complexity, and security of privacy-preserving string comparators for record linkage. *Information Fusion, 13*(4), 245-259.

Earley, P. C., & Kanfer, R. (1985). The influence of component participation and role models on goal acceptance, goal satisfaction and performance. *Organizational Behavior and Human Decision Processes, 36*(3), 378-390.

Ferrante, A., & Boyd, K. (2012). A transparent and transportable methodology for evaluating Data Linkage software. *Journal of Biomedical Informatics, 45*(1), 165-172.

Fox, M. S., Barbuceanu, M., Gruninger, M., & Lin, J. (1998). An organization ontology for enterprise modelling *Simulation organizations: Computational models of institutions and groupsAAAI/MIT Press* (pp. 131-152).

Freire, S. M., Almeida, R. T. d., Cabral, M. D. B., Bastos, E. d. A., Souza, R. C., & Silva, M. G. P. d. (2012). A record linkage process of a cervical cancer screening database. *Computer Method and Program in Biomedecine, 108*(1), 90-101.

Herzog, T. N., Sheuren, F. J., & Winkler, W. E. (2007). *Data quality and record linkage technique*. Washington, USA: Spinger.

Izhar, T. A. T., Torabi, T., Bhatti, I., & Liu, F. (2012). *Analytical dependency between organisational goals and actions: Modelling concept.* Paper presented at the International Conference on Innovation and Information Management (ICIIM 2012) Chengdu, China.

Izhar, T. A. T., Torabi, T., Bhatti, M. I., & Liu, F. (2013). Recent developments in the organization goals conformance using ontology. *Expert Systems with Applications, 40*(10), 4252-4267.

Jutte, D. P., Roos, L. L., & Brownell, M. D. (2011). Administrative record linkage as a tool for public health research. *Annual Review Public Health, 32*, 91-108.

Kang, S., & Norton, H. E. (2004). Nonprofit organizations' use of the World

Wide Web: are they sufficiently fulfilling organizational goals. *Public Relations Review, 30*(3), 279-284.

Karakasidis, A., & Verykios, V. S. (2011). Secure blocking+secure matching= secure record linkage. *Journal of Computing Science and Engineering, 5*(3), 223-235.

Kum, H.-C., Duncan, D. F., & Stewart, C. J. (2009). Supporting self-evaluation in local government via Knowledge Discovery and Data Mining. *Government Information Quarterly, 26*(2), 295-304.

Lepmets, M., McBride, T., & Ras, E. (2012). Goal alignment in process improvement. *The Journal of System and Software, 85*(6), 1440-1452.

Liao, S.-H., Chang, W.-J., & Lee, C.-C. (2008). Mining marketing maps for business alliances. *Expert Systems with Applications, 35*(3), 1338-1350.

Mansingh, G., Osei-Bryson, K.-M., & Reichgelt, H. (2009). Building ontology-based knowledge maps to assist knowledge process outsourcing decisions. *Knowledge Management Research and Practice, 7*, 37-51.

Meray, N., Reitsma, J. B., Ravelli, A. C. J., & Bonsel, G. J. (2007). Probabilistic record linkage is a valid and transparent tool to combine databased without a patient identification number. *Journal of Clinical Epideminology, 60*(9), 883-891.

Mikroyannidis, A., & Theodoulidis, B. (2010). Ontology management and evolution for businee intelligence. *International Journal of Information Management, 30*(6), 559-566.

Pundt, H., & Bishr, Y. (2002). Domain ontologies for data sharing-an example from environmental monitoring using field GIS. *Computer & Geosciences, 28*(1), 95-102.

Rao, L., Mansingh, G., & Osei-Bryson, K.-M. (2012). Building ontology based knowledge maps to assist business process re-engineering. *Decision Support Systems, 52*(3), 577-589.

Romero, O., & Abello, A. (2010). A framework for multidimensional design of data warehouses from ontologies. *Data & Knowledge Engineering, 69*(11), 1138-1157.

Salerno, M. S. (2009). Reconfigurable organisation to cope with unpredictable goals. *International Journal Economics, 122*(1), 419-428.

Scannapieco, M., Figotin, I., Bertino, E., & Elmagarmid, A. (2007, 12-14 June). *Privacy preserving schema and data matching.* Paper presented at the ACM SIGMOD Conference, Beijing, China.

Selma, K., Ilyes, B., Ladjel, B., Eric, S., Stephane, J., & Michael, B.

(2012). Ontology-based structured web data warehouses for sustainable interoperability: requirement modeling, design methodology and tool. *Computer in Industry, 63*(8), 799-812.

Sharma, S., & Osei-Bryson, K.-M. (2008). *Organization-ontology based framework for implementing the business understanding phase of data mining projects.* Paper presented at the International Conference on System Sciences, Hawaii.

Sholihin, M., Pike, R., Mangena, M., & Li, J. (2011). Goal-setting participation and goal commitment: Examining the mediating roles of procedural fairness and interpersonal trust in a UK financial services organisation. *The British Accounting Review, 43*(2), 135-146.

Su, W., Wang, J., & Lochovsky, F. H. (2010). Record matching over query results from multiple web databases. *IEEE Transaction on Knowledge and Data Engineering, 22*(4), 578-589.

Trepetin, S. (2008). Privacy-preserving string comparisons in record linkage systems: A review. *Information Security Journal: A Global Perspective, 17*(5-6), 253-266.

Varghese, C. E., & Sundar, G. N. (2011). Record matching: Improving performance in classification. *International Journal on Computer Science and Engineering, 3*(3), 1207-1212.

Yakout, M., Elmagarmid, A. K., Elmeleegy, H., & Ouzzanil, M. (2010). *Behavior based record linkage.* Paper presented at the The 36th International Conference on Very Large Data Bases, Singapore.