

Utilising ontology for “heterogeneous data analysis in organizational goals”

Tengku Adil Tengku Izhar

Universiti Teknologi MARA, Puncak Perdana Campus, Shah Alam, Malaysia

Bernady O. Apduhan

Kyushu Sangyo University, Fukuoka, Japan, and

Torab Torabi

La Trobe University, Melbourne, Australia

Abstract

Purpose – The purpose of this paper is to assess the level of the organizational goal accomplishment by assessing the reliance relationship between organizational data and organizational goals.

Design/methodology/approach – The evaluation of the organizational goals is based on design and operational level, which can serve in ranking of the organizational goals achievement and hence assist the decision-making process in achieving the organizational goals. To achieve this aim, the authors propose an ontology to develop the relationship between organizational data and organizational goals.

Findings – 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 the authors suggest that these data attributes are relevant in relation to the organizational goals.

Originality/value – The contribution of this paper will serve as the first step to evaluate the relevance of organizational data to assist decision-making in relation to the organizational goals.

Keywords Ontology, Data linkage, Dataset, Organizational goals ontology, Organizational goals

Paper type Research paper

1. Introduction

Government agencies, public and private bodies are suffocating in a steadily expanding storm of data as they create huge measure of data in their day by day business movement (Christen, 2012). The capacity to break down their data in a timely manner can give an aggressive edge to enhance profitability in connection to the organizational goals. Data are the most critical advantage for help the decision-making process in accomplishing the organizational goals. In any case, the reliability of organizational data in connection to the organizational goals is constantly faulty because of the gigantic data that have been made inside the organization. Some of this data get to be not applicable or after some time get to be unessential to the organization goals. Despite the fact that professionals such as data analysts are trained to analyze these data but the increase amount of organizational data has become a major problem in applying these data for achieving the organizational goals. The growth in the amount of the organization resources available nowadays poses major difficulties and challenges in decision-making (Mikroyannidis and Theodoulidis, 2010). The challenge is to identify the relevance of organizational data. Therefore,



demonstrating the organizational goals structure is essential to build up a relationship between authoritative information and hierarchical objectives. For example, an ontology improves the understanding of the organizational goals structure as it shows the relationship of the organizational goal elements (Izhar *et al.*, 2012; Izhar *et al.*, 2013).

Most studies which have been conducted in this issue focus on business intelligent (Nofal and Yusof, 2013; Azma and Mostafapour, 2012; Popovic *et al.*, 2012), data mining (Weerd *et al.*, 2013; Zandi, 2014; Aghdaie *et al.*, 2014), data linkage (Christen, 2008; Durham *et al.*, 2012; Ferrante and Boyd, 2012) and knowledge discovery in databases (Lin *et al.*, 2008; Cheng *et al.*, 2012). Even though these studies focus on decision-making but they do not focus on the interaction between organizational data and organizational goals, as shown in Figure 1. As a result, it is difficult to identify the relevance of organizational data that relate to the organizational goals. We tackle the problem to evaluate the organizational goals by proposing the organizational goals model based on an ontology to assist the decision-making process in relation to the organizational goals. We propose this model to evaluate organizational data to support decision-making and thereby assist the organization to achieve its goals. The model also important in measuring the extent to which organizational data are consistent with organizational goals (Izhar *et al.*, 2013). To develop this organizational goals model, we observe the limitations from the previous studies on the organizational model.

The aim of this paper is to develop the organizational goals ontology to evaluate the dependency relationship between organizational data and organizational goals. The main challenge is to identify which organizational data are relevant from organizational data sets that relate to the organizational goals. In this paper, the evaluation of the organizational goals ontology is based on design and operational level, in which can serve in ranking of the organizational goals achievement and hence assist the decision-making process in achieving the organizational goals.

The remainder of this paper is organized as follows. Section 2 discusses on organizational goals ontology. Evaluation level of the organizational goals ontology is discusses in Section 3. Section 4 is a case study. Discussion and future research are discusses in Section 5. Section 6 is a conclusion.

2. Organizational goals ontology

The main contribution of this section is to address the issues debate on the organizational resources respectively and takes into account both the evaluation of the organizational data and the organizational goals. To our knowledge, there is no study consider the evaluation of organizational data to assist decision-making process in relation to the organizational goals. Therefore, it is important to identify the linkage between organizational data and organizational goals (3). While many studies have been carried out on data linkage in wide areas such as data privacy, medical data, software and databases (13, 28, 29), limited study has been observed in evaluating organization data that relate to the organizational goals (3). Therefore, the first step to identify the relevance of organizational data is to recognize the relationship of organizational data that relate the organizational goals. In contrast to the previous studies on data linkage (13, 28, 29), we adapted data linkage to address the

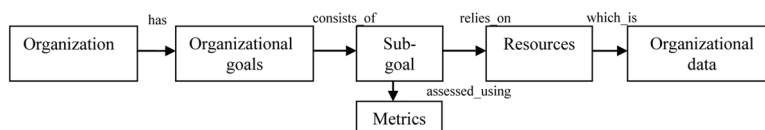


Figure 1.
Organizational goals
ontology

dependency relationship between organizational data and organizational goals. This is because organizational data are a major resource in every organization and it is important to identify the relevance of this data for the achievement of the organizational goals. It is the 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 (38, 39), goal performance (34, 37, 40) and organization performance (32, 33, 35, 36). Therefore, the linkage of organizational data from data sets should be consistent to prevent any redundant data so it can be evaluated in relation to the organizational goals.

Instead of evaluating the information and knowledge, we argue that the organizational data are important and must be considered during the decision-making process in relation to the organizational goals because it is very important to identify the relevance of organizational data to achieve the organizational goals.

Table I shows the results from the previous models on the organizational goals using an ontology but none of these studies focus on the evaluation level of the organizational goals achievement. Table I also shows that these models do not focus on organizational data. There is no study on metrics to evaluate the dependency relationship of organizational data that relate to the organizational goals. The gaps of these issues are important during the development of the organizational goals ontology.

Compare to Sharma and Osei-Bryson (2008), our organizational goals ontology focus on the usage of organizational data instead of knowledge, information or tools because organizational data are a major resource in every organization and it is important to evaluate the relevance of this organizational data in achieving the organizational goals. We also suggest organizational data are important as information and knowledge to assist the decision-making process (Izhar *et al.*, 2013).

In organization, it is extremely important for the manager to have access to the most relevant organizational data in relation to the organizational goals. Pointed out that sharing important data and information can provide the required knowledge to assist successful decision-making. It is crucial for organizations to create and generate new data and evaluate it to enhance decision-making. Different ways of generating new ideas, information and knowledge will help in terms of decision-making and will enable teams within the organization to use the most relevant organizational data to successfully achieve the organizational goals.

In this paper, the first step to identify the relevant organizational data is to recognize the matching set of organizational data to identify which organizational data relate to the organizational goals. A large body of existing work uses the terms record linkage, data linkage, record matching, data matching, we use the term data dependency in an effort to identify the dependency relationship between the organizational data and organizational

Authors	Organizational goals ontology		Resources		Metrics	
	Organizational goals	Sub-goals	Information	Knowledge	Organizational data	Dependency relationship
Fox <i>et al.</i> (1998)	/	/	×	×	×	×
Sharma and Osei-Bryson (2008)	/	/	/	/	×	×
Rao <i>et al.</i> (2012)	/	/	/	/	×	×

Table I.
Review of the issues

goals because we attempt to identify dependency for all organizational data that relate to organizational goals (Figure 2).

2.1 Metrics definition

The weight to analyze the dependencies of organizational data can be defined in many ways, such as percentage and frequency based on different situations. For example, domain experts and entrepreneurs might want to identify the percentage of data that relate to organizational goals. After we identify this value, it can be presented on the dashboard to show a graphical presentation of value. The comparison of this value can be presented to support the decision-making process in relation to the organizational goals.

Varghese and Sundar (2011) evaluated the weight of the matching data for metrics evaluation that used various types of databases (Herzog *et al.*, 2007). For example, we assumed two organizational data, data a and data b , from two organizational data sets, data set A and data set B , relate to the same organizational goals, $a \in A$ and $b \in B$. For example, if both organizational data a and b relate to the same organizational goals, then

$$M = \{(a, b) : a \in A, b \in B, (a, b)\}$$

Is matching organizational data between data a and b to the same organizational goals.

In this paper, how domain experts and entrepreneurs want to define their own metrics is not the main objective of the paper. This is because they might want to evaluate the organizational data and define the metrics in different ways.

2.2 Frequency value

In this paper, we define the metrics based on the frequency value to evaluate the weight for the dependency relationship between organizational data and organizational goals. In contrast to Ferrante *et al.* (Ferrante and Boyd, 2012), the authors did not define the weight that be used in linkage of data. Frequency value in this example can be calculated using the following metrics:

$$\text{Frequency value} = \left(\frac{\text{Number}}{\text{Total number}} \right)$$

In Christen (Christen, 2008), the author discuss 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 *et al.*, 2008). In this paper, 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, $1 \geq a \geq 0$.



Figure 2.
Data goals
dependency

3. Case study

The case study used in this paper aims to evaluate the level of student satisfaction in the La Trobe Student Support Services at La Trobe University, Melbourne Australia. The aim of the La Trobe University Support Services is to improve the students' university experience by providing services that encourage students to socialize and become involved in things other than academic activities.

In this case study, a value is assigned to each La Trobe Student Support Service to identify the degree of student satisfaction of the student support services and to identify which service was considered the most important in increasing student satisfaction.

In this case study, we applied data set from La Trobe University Student Support Services Experiment Report. In this data set, we select ten different services provided by La Trobe University based on two different student backgrounds. It is important to note that we do not mean to imply that the entire services are not important but we suggest that these numbers of services with only two student backgrounds are enough to test the flexibility of the organizational goals ontology with respect to the La Trobe University Student Support Services.

3.1 Identify the goal

The organizational goals ontology identifies the relationship for the case study goal. The experiment on the organizational goals ontology addressed the relationship of two different variables evaluated to identify student satisfaction to the La Trobe Student Support Services. Therefore, the scope of goal definition for the case study presented as:

- Goal 1: Level of student satisfaction for student who have not used the services;
- Goal 2: Level of student satisfaction for student who have used the services; and
- Variables: Student background.

To evaluate the level of the student satisfaction, data are evaluated based on student background. There are two categories of student background are selected from the La Trobe University Student Support Services Experiment Report:

- (1) Level of study consists of postgraduate and undergraduate; and
- (2) Campus consists of Bundoora, Bendigo, Mildura, Shepparton, Albury Wodonga and City (Figure 3).

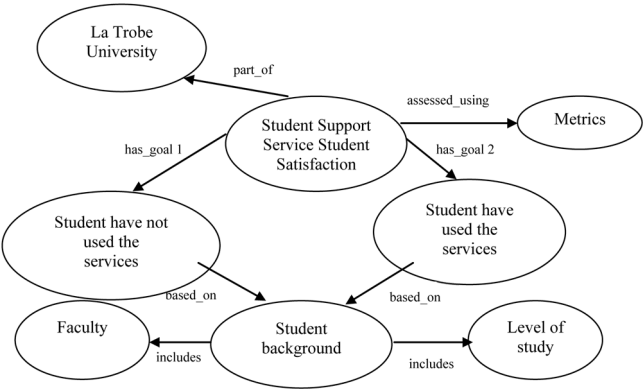


Figure 3.
Ontology for La
Trobe university
student support
services

3.2 Identify the relationship between data and goal

In contrast to Christen (Christen, 2012), we elaborate the process of data linkage in the context of the organizational goals, as shown in Figure 4. In this figure, we show the elements of data linkage (Christen, 2012) and we consider these elements as terms used in the previous studies such as data sets (Durham *et al.*, 2012; Su *et al.*, 2010), data pre-processing (Trepetin, 2008), data linkage (Christen, 2012), data matching (Su *et al.*, 2010; Christen, 2012). However, previous discussion covered the linkage for domain data such as systems and databases (Trepetin, 2008; Ferrante and Boyd, 2012). To our knowledge, there is no discussion cover the element of organizational data. Past studies observed data linkage as a task of identifying data corresponding to the same entity from one or more data sources (Christen, 2012). Previous studies only evaluate data that match to the same entities such as databases. In respond to this gap, we observe the existing data attributes from the organizational data sets to evaluate the value of each existing data that relate to the organizational goals.

3.2.1 Data sets. Data sets are the collection of data that have been stored in different set of storage that represent different meaning of data. These data sets are important as a reference for any decision-making evaluation. However, these data can be very large and they create a challenge on how to identify relevance of data from huge number of data sets. Therefore, it is important to discuss the data sets creation to perform any evaluation of data from database (Durham *et al.*, 2012; Su *et al.*, 2010). In contrast to our work, we attempt to perform the evaluation of organizational data in relation to the organizational goals. The evaluation is important to identify relevance of organizational data from vast amount of data sets that relate to the organizational goals. The process is important to assist domain experts or decision makers in making any decision in relation to the organizational goals.

3.2.2 Data pre-processing. Organizations create vast amount of data every day, and some of these data are not relevant for decision-making (Izhar *et al.*, 2012). Christen (Christen, 2012) defined data pre-processing as a process to identify quality data. In contrast, we suggest data pre-process is important to identify the existing data in relation to the organizational goals. At the same time, it is important to identify matching data from various data sets (Trepetin, 2008).

3.2.3 Data attributes. In this paper, we suggest the process of identifying data attributes from data sets is important as an effort to identify the relevance of organizational data. In contrast to the previous studies (Trepetin, 2008; Ferrante and Boyd, 2012), the authors observed the linkage of domain data such as databases and there is no observation has been carried out in the context of organizational data in relation to the organizational goals. Data attribute identification is important to develop the dependency relationship between organizational data that relate to the organizational goals. For example, let's take La Trobe University Student Support Service Survey as an example. The Division of Student Services aims to improve the university experience by providing services that encourage students to socialize and become involved in things other than books.

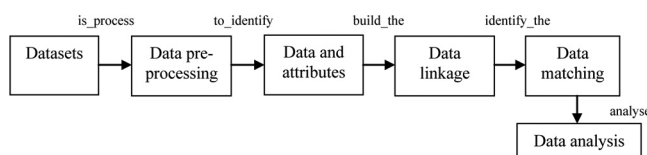


Figure 4.
Data dependency
process to the
organizational goals
(adapted from
Christen, 2012)

In this example, the survey looks at “what student believe is important if they have not used the university service” and “what student believe is important if they have used the university service”. Therefore, data sets for these questions are based on the university services as shown in [Tables II](#) and [III](#).

3.2.4 Data linkage. In this paper, we refer data linkage as a task of finding data in data sets that refer to the same organizational goals. We suggest data linkage is important to identify existing data from large data sets during the decision-making process in relation to the organizational goals. At the same time, data linkage is important in joining data sets based on entities that may or may not share the same organizational goals. However, we use the term data dependency to show the relationship of organizational data that relate to the organizational goals.

3.2.5 Data matching. Data matching is the important step in identifying data from data sets ([Su et al., 2010](#); [Christen, 2012](#)). We refer data matching as data from data sets that relate to the organizational goals. We suggest data matching is a process of bringing together data from different sources that relate to the organizational goals. In contrast, we attempt to identify the value of this data to consider these data are relevant in relation to the organizational goals. Data matching can be beneficial to the organizational goals because

Table II.
Data set for “what student believe is important if they have not used the service

No.	What student believe is important if they have not used the service	Data attributes
1	University scholarship	University, scholarship
2	Career planning and advice	Career, planning, advice
3	Career information and resources	Career, information, resources
4	Work integrated learning (co-op, placements, projects)	Work integrated learning, co-op, placements, projects
5	Learning support- outside class	Learning support, outside class
6	English language support	Support, English language
7	Leadership skills development	Leadership, skills development
8	Discrimination and harassment support services	Support, services, discrimination, harassment
9	Student legal service	Service, student legal
10	Financial advice and student loans	Financial, advice, student loans

Table III.
Data set for “what student believe is important if they have used the service

No.	What student believe is important if they have used the service	Data attributes
1	University scholarship	University, scholarship
2	Career planning and advice	Career, planning, advice
3	Career information and resources	Career, information, resources
4	Work integrated learning (co-op, placements, projects)	Work integrated learning, co-op, placements, projects
5	Sporting activities/ facilities	Sporting, activities, facilities
6	English language support	Support, English language
7	Student portal announcement and news	Student portal, announcement, news
8	Disability support	Support, disability
9	La Trobe's Social media (Twitter, Facebook)	La Trobe's Social media, Twitter, Facebook
10	Chaplaincy and religious services	Services, chaplaincy, religious

they enable to determine the correct value and reducing the errors of data that do not relate to the organizational goals.

3.3 Analysis and results

In this paper, we measure the weight of data based on the frequency value. First, we identify the frequency value of two different data sets in relation to goal 1 and goal 2 based on postgraduate and undergraduate student perspective as shown in [Tables IV](#) and [V](#). [Tables IV](#) and [V](#) show the frequency value of each data that match to goal 1 and goal 2. The results show partial support for the proposed concept to identify the frequency value for each data to identify highest value of data to be considered relevant.

The results in [Tables IV](#) and [V](#) conclude that undergraduate students believe that university services are important to them. We reach to the decision that the respondent from undergraduate students believe that the services provided by the university is important rather if they use the services or not. The results also show that the number of undergraduate students is higher than postgraduate students. Therefore, the value should be higher for undergraduate students. In these tables, the process of decision-making is not

Goal 1 Level of study/ Services	What student believe is important if they have not used the service			
	Postgraduate		Undergraduate	
	No.	Frequency value	No.	Frequency value
University scholarship	490	0.172111	2,357	0.827889
Career planning and advice	529	0.186399	2,309	0.813601
Career information and resources	496	0.194586	2,053	0.805414
Work integrated learning (co-op, placements, projects)	500	0.196618	2,043	0.803382
Learning support- outside class	375	0.199045	1,509	0.800955
English language support	542	0.178172	2,500	0.821828
Leadership skills development	589	0.194069	2,446	0.805931
Discrimination and harassment support services	633	0.191065	2,680	0.808935
Student legal service	619	0.188375	2,667	0.811625
Financial advice and student loans	616	0.193407	2,569	0.806593

Table IV.
Frequency value of
“what student
believe is important
if they have not used
the service”

Goal 2 Level of study/ Services	What student believe is important if they have used the service			
	Postgraduate		Undergraduate	
	No.	Frequency value	No.	Frequency value
University scholarship	153	0.294798	366	0.705202
Career planning and advice	94	0.207506	359	0.792494
Career information and resources	125	0.171233	605	0.828767
Work integrated learning (co-op, placements, projects)	123	0.166216	617	0.833784
Sporting activities/ facilities	175	0.168269	865	0.831731
English language support	78	0.369668	133	0.630332
Student portal announcement and news	275	0.17209	1323	0.82791
Disability support	23	0.163121	118	0.836879
La Trobe's Social media (Twitter, Facebook)	152	0.125932	1055	0.874068
Chaplaincy and religious services	39	0.208556	148	0.791444

Table V.
Frequency value of
“what student
believe is important
if they have used the
service”

complicated because it based on two elements; postgraduate students and undergraduate students (Figures 5 and 6).

Second, we identify the frequency value of data in relation to goal 1 and goal 2 based on the faculty. We interpret the results to identify the frequency value of data for every faculty as shown in Tables VI and VII (Figures 7 and 8).

In Tables VI and VII, we identify the frequency value for each data to identify data to be considered relevant. In these tables, we can see the comparison of the value for every faculty. For example, the results show that *university scholarship* has the highest value for the Health Science students for goal 1 in Table VI. Therefore, we make a decision that students from the Faculty of Health Science believe that *university scholarship* is the most important service provided by the university even though the student do not receive the scholarship. The results in Table VII also show that Health Science students believe that *university scholarship* is important to them compare to the other university services. We can make a decision that *university scholarship* has the highest value from the Health Science students for goal 1 and goal 2 compare to the other students from other faculties.

To support our explanation, let's take *university scholarship* as an example. In Table VI and Table VII, the results show that *university scholarship* has the highest frequency value as shown in Figure 9. Students from the Faculty of Health Science believes that *university scholarship* is the most important service provided by La Trobe University whether they use the service or not. We can make a decision that La Trobe University provides variety of scholarship with an industry scholarship program for student to find more information about *university scholarship* as the scholarship cover the range of undergraduate students

Figure 5.
Summary of
frequency value for
Goal 1 based on study
level

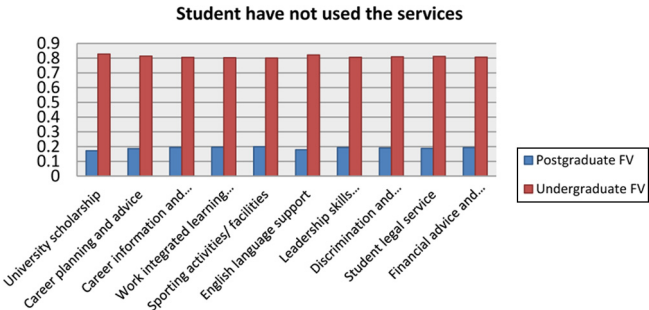
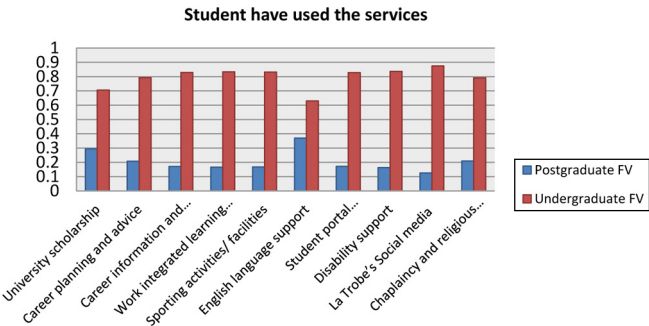


Figure 6.
Summary of
frequency value for
Goal 2 based on study
level



Goal 1 Faculty/ Services	What student believe is important if they have not used the service									
	Science, technology and engineering		Law and management		Humanities and social science		Health science		Education	
	N	FV	N	FV	N	FV	N	FV	N	FV
University scholarship	496	0.175079	580	0.20473	554	0.195552	909	0.320861	294	0.103777
Career planning and advice	532	0.188185	500	0.176866	549	0.194199	941	0.332862	305	0.107888
Career information and resources	462	0.181818	482	0.189689	546	0.214876	804	0.316411	247	0.097206
Work integrated learning (co-op, placements, projects)	527	0.207808	566	0.223186	589	0.232256	662	0.261041	192	0.07571
Learning support- outside class	371	0.197551	369	0.196486	405	0.215655	567	0.301917	166	0.088392
English language support	574	0.189501	561	0.18521	598	0.197425	977	0.322549	319	0.105315
Leadership skills development	554	0.183322	600	0.198544	587	0.194242	982	0.32495	299	0.098941
Discrimination and harassment support services	624	0.189091	660	0.2	633	0.191818	1,043	0.316061	340	0.10303
Student legal service	614	0.187595	648	0.197984	628	0.191873	1,040	0.317751	343	0.104797
Financial advice and student loans	589	0.185804	637	0.200946	609	0.192114	1,007	0.3176	328	0.10347

Notes: N = Number; FV = Frequency value

Table VI.
Frequency value of
“what student
believe is important
if they have not used
the service” based on
the faculty

IJWIS
15,3

Goal 2 Faculty/ Services	What student believe is important if they have used the service											
	Science, technology and engineering				Law and management		Humanities and social science		Health science		Education	
	N	FV	N	FV	N	FV	N	FV	N	FV	N	FV
University scholarship	139	0.267823	96	0.184971	87	0.16763	146	0.28131	51	0.098266		
Career planning and advice	91	0.202673	149	0.331849	79	0.175947	98	0.218263	32	0.071269		
Career information and resources	159	0.219613	165	0.227901	77	0.106354	233	0.321823	90	0.124309		
Work integrated learning (co-op, placements, projects)	94	0.128415	81	0.110656	38	0.051913	373	0.509563	146	0.194454		
Sporting activities/facilities	230	0.222222	217	0.209662	144	0.13913	337	0.325604	107	0.103382		
English language support	37	0.174528	83	0.391509	24	0.113208	51	0.240566	17	0.080189		
Student portal announcement and news	278	0.174295	365	0.22884	321	0.201254	484	0.303448	147	0.092163		
Disability support	40	0.28169	22	0.15493	36	0.253521	32	0.225352	12	0.084507		
La Trobe's Social media (Twitter, Facebook)	238	0.198003	271	0.225458	224	0.186356	348	0.289517	121	0.100666		
Chaplaincy and religious services	49	0.264865	33	0.178378	33	0.178378	50	0.27027	20	0.108108		

Notes: N = Number; FV = Frequency value

and postgraduate students and also for international students. The evaluation results show that *university scholarship* program can support the goal to increase the number of students as part of La Trobe Strategy Plan in relation to the university's goal.

4. Discussion

We identify the frequency value for data in relation to the goal 1 and goal 2. Take *university scholarship* in Tables VI and VII as an example. In these tables, we can see the frequency value for *university scholarship*. The results based on the faculty show that student from the Faculty of the Health Science have the highest value compare to other faculties for both goals. However, Faculty of Law and Management have the second highest value for goal 1 but has the third highest value for goal 2. We can see different results for Humanities and

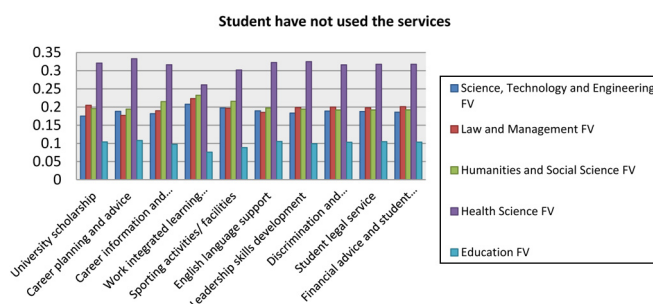


Figure 7.
Summary of
frequency value for
Goal 1 based on the
faculty

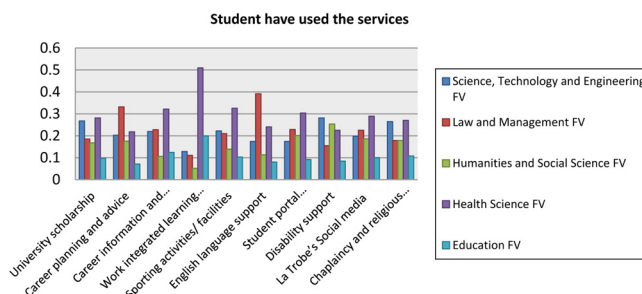


Figure 8.
Summary of
frequency value for
Goal 2 based on the
faculty

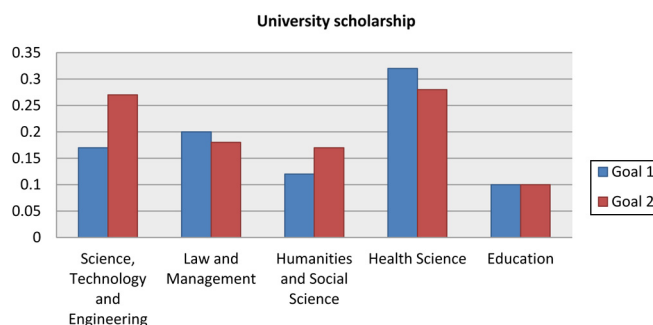


Figure 9.
Frequency value of
the university
scholarship for Goal 1
and Goal 2 based on
the faculty

Social Science, and Science, Technology and Engineering. We conclude that frequency value assist the decision-making process in relation to the La Trobe University Student Support Service Survey's goals. We conclude that students consider *university scholarship* is the most important service provided by La Trobe University. At the same time, the results show that university scholarship program can support the La Trobe University's goals in increasing the number of students as part of the La Trobe University Strategy Plan.

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.

5. Conclusion

In this paper, the organizational goals ontology provide a two-way approach:

- (1) to identify which organizational goals have been achieved; and
- (2) to identify which organizational data are more important to the organizational goals.

In the future, not only will this model be able to identify which organizational goals are more important and achievable, it will have the capability to identify the organizational data first. In the future, we aim to incorporate social data in the organizational goals ontology to identify how social data can be applied to evaluate the degree to which the organizational goals have been achieved. Therefore, our organizational goals ontology could capture this data and develop the dependency relationship between this data and the organizational goals. Due to the flexibility of the organizational goals ontology, it is able to capture large sets of data that relate to the organizational goals that enable this data to be analyzed and the results of the analyzed data can assist the decision-making process in relation to the organizational goals.

References

- Aghdaie, M.H., Zolfani, S.H. and Zavadskas, E.K. (2014), "Synergies of data mining and multiple attribute decision making", *Procedia-Social and Behavioral Sciences*, Vol. 110, pp. 767-776.
- Arasu, A., Chaudhuri, S. and Kaushik, R. (2008), "Transformation-based framework for record matching", *IEEE 24th International Conference, Cancun*, pp. 40-49.
- Azma, F. and Mostafapour, M.A. (2012), "Business intelligence as a key strategy for development organizations", *Procedia Technology*, Vol. 1, pp. 102-106.
- Cheng, T., Wang, Y. and Sun, Y. (2012), "Development and application of tender evaluation decision-making and risk early warning system for water projects based on KDD", *Advances in Engineering Software*, Vol. 48, pp. 58-69.
- Christen, P. (2008), "Automatic record linkage using seeded nearest neighbour and support vector machine classification", *14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Las Vegas, NV*, pp. 151-159.
- Christen, P. (2012), "A survey of indexing techniques for scalable record linkage and deduplication", *IEEE Transactions on Knowledge and Data Engineering*, Vol. 24 No. 9, pp. 1537-1555.
- Durham, E., Xue, Y. and Kantarcioglu, M. (2012), "Quantifying the correctness, computational complexity, and security of privacy-preserving string comparators for record linkage", *An International Journal on Information Fusion*, Vol. 13 No. 4, pp. 245-259.

- Ferrante, A. and Boyd, K. (2012), "A transparent and transportable methodology for evaluating data linkage software", *Journal of Biomedical Informatics*, Vol. 45 No. 1, pp. 165-172.
- Fox, M.S., Barbuceanu, M. and Gruninger, M. (1998), "An organization ontology for enterprise modelling", *Simulation Organizations: Computational Models of Institutions and Groups*, AAAI/MIT Press, Massachusetts, pp. 131-152.
- Herzog, T.N., Sheuren, F.J. and Winkler, W.E. (2007), *Data Quality and Record Linkage Technique*, Springer, Washington, DC.
- Izhar, T.A.T., Torabi, T. and Bhatti, I. (2012), "Analytical dependency between organisational goals and actions: modelling concept", *International Conference on Innovation and Information Management (ICIIM 2012)*, IACSIT Press, Chengdu.
- Izhar, T.A.T., Torabi, T. and Bhatti, M.I. (2013), "Recent developments in the organization goals conformance using ontology", *Expert Systems with Applications*, Vol. 40 No. 10, pp. 4252-4267.
- Lin, C., Lin, C.M. and Li, S.T. (2008), "Intelligent physician segmentation and management based on KDD approach", *Expert Systems with Applications*, Vol. 34 No. 3, pp. 1963-1973.
- Mikroyannidis, A. and Theodoulidis, B. (2010), "Ontology management and evolution for business intelligence", *International Journal of Information Management*, Vol. 30 No. 6, pp. 559-566.
- Nofal, M.I. and Yusof, Z.M. (2013), "Integration of business intelligence and enterprise resource planning within organizations", *Procedia Technology*, Vol. 11, pp. 658-665.
- Popovic, A., Hackney, R. and Coelho, P.S. (2012), "Towards business intelligence systems success: effects of maturity and culture on analytical decision making", *Decision Support Systems*, Vol. 54 No. 1, pp. 729-739.
- Rao, L., Mansingh, G. and Osei-Bryson, K.-M. (2012), "Building ontology based knowledge maps to assist business process re-engineering", *Decision Support Systems*, Vol. 52 No. 3, pp. 577-589.
- Sharma, S. and Osei-Bryson, K.M. (2008), "Organization-ontology based framework for implementing the business understanding phase of data mining projects", *International Conference on System Sciences, HI*, p. 27.
- Su, W., Wang, J. and Lochovsky, F.H. (2010), "Record matching over query results from multiple web databases", *IEEE Transactions on Knowledge and Data Engineering*, Vol. 22 No. 4, pp. 578-589.
- Trepetin, S. (2008), "Privacy-preserving string comparisons in record linkage systems: a review", *Information Security Journal: A Global Perspective*, Vol. 17 Nos 5/6, pp. 253-266.
- Varghese, C.E. and Sundar, G.N. (2011), "Record matching: improving performance in classification", *International Journal on Computer Science and Engineering*, Vol. 3, pp. 1207-1212.
- Weerdt, J.D., Schupp, A. and Vanderloock, A. (2013), "Process mining for the multi-faceted analysis of business processes- A case study in a financial services organization", *Computers in Industry*, Vol. 64 No. 1, pp. 57-67.
- Zandi, F. (2014), "A bi-level interactive decision support framework to identify data mining- oriented electronic health record architectures", *Applied Soft Computing*, Vol. 18, pp. 136-145.

Corresponding author

Tengku Adil Tengku Izhar can be contacted at: tengkuadil@yahoo.co.uk

Reproduced with permission of copyright owner. Further reproduction prohibited without permission.