International Journal of Computer Information Systems and Industrial Management Applications.
ISSN 2150-7988 Volume 10 (2018) pp. 001-007

Cloud Based Enterprise Global Ontology For Information Enterprise: A Proposed Framework

Tengku Adil Tengku Izhar[1] and Bernady O. Apduhan[2]

[1] Faculty of Information Management
Universiti Teknologi MARA, UiTM
Shah Alam, Selangor, Malaysia
tengkuadil4540@salam.uitm.edu.my

[2] Faculty of Information Science
Kyushu Sangyo University,
Fukuoka, Japan
boa93dai@gmail.com

Abstract: Big Data in organizations have transformed the way organizations across industries implement new approach to handle huge amount of data. Organizations rely to this data to achieve specific business priorities. The challenge is how to capture this data to be considered relevant for the specific organization activities because determining relevant data is a key to deliver information from massive amounts of data. The aim of this paper is to integrate external data using an ontology to capture relevant information for efficient decision-making. In order to achieve this aim, we tackle the literature to incorporate cloud and ontology to retrieve external data such as social media and internal data such as organizational data. The results can lead to some new evaluation methods in big data era from different perspectives. The outcome will offer an enormous opportunity to advance the science of data analytics so that future researchers will have a new understanding on what is needed to improve their data analysis process. The research benefit nation, economy and society. The conduct of this survey will ensure the projects agility in responding to unfolding events, and substantially enhance its ability to engage in and impact on organizations and societies.

Keywords: big data, cloud, enterprise, information retrieval, ontology

I. Introduction

Big data provides significant opportunities for enterprises to impact a wide range of business processes in organizations [1-4]. Although there are many studies conducted on Big Data in the context of the organizations [2, 5-7], there is still little debate these days on the role and importance of big data for efficient decision-making [7-9]. Companies like Google, eBay, LinkedIn, and Facebook were built around big data from the beginning [8]. There is yet no consensus on how best to incorporate big data in organizations and how the process of incorporating big data can identify the relevance of data to assist in decision-making process. The emergence of big data has led to a profound transformation on how organizations store, manage and analyze their data. Increased access to large-scale data enables an organization to capture relevant data, which can assist in their decision-making process [1]. It is now possible to manage the volume, velocity and variety of data using the tools and software, so the focus in this research is to shift towards how we can use analyses to find extra value from this data analysis. This will make significant achievement that provides an important innovation in research methods in big data era to trace how data flows across organizations. Furthermore, it will educate organizations on how to analyze data to support their decision-making process.

Even though data scientists are trained to analyze data, but with the huge volume of data generated everyday makes it harder to identify which data is relevant to the organization's specific activity. As a result, it poses an issue on how to effectively utilize this data to support decision-making processes [1, 10-14]. Big data will be important to business organizations because more data can create more accurate analysis which can in turn lead to more confident decision-making. Taking advantage of big data opportunities is a challenge for organizations [2]. In order to ensure the effectiveness of the data, organizations need to store the data reliably across a number of databases. Once the data is distributed, and when the needs arise, the organization must find a way to extract the data again, identify which data is needed, assemble it and analyze it. The challenge now is how to capture relevant data from this massive amount of data, which can deliver values to the organization's specific activities. The real issue is not how the organizations will acquire the huge volume of data, but how they can harness the value of this data that counts [8]. Data is critical for the survival and growth of organizations and people. The challenges for domain experts are now less about managing activities that collect, store and disseminate data. There is greater focus on managing activities that make changes in patterns of behaviour of customers, people, and organizations, and information that leads to changes in the way organization use data to engage in knowledge focused activities.

The aim of this research is to develop a framework that underpins a seamless integration of organizational data and heterogeneous external data pertinent to the organization's focus area. In this era of big data science, it is critical for organizations and businesses to be able to embrace this new facility and to accurately integrate the knowledge-bases from multiple sources into the organizational information repository. The availability of a mechanism that allows seamless consolidation of knowledge from external sources will enrich the capability of the organization to make accurate decision-making. These heterogeneous external sources are growing very significantly in the last few years, especially due to the availability of wireless and mobile technologies, crowd-sourcing facilities, Internet of Things and sensor networks, as well as social media and web data. All these technologies generate huge amount of data and together they can be extracted to generate values to the organization and to establish situational awareness of the community or market trends.

II. Literature Review

Organizations have at their disposal a large volume of data with a wide variety of types [15-24]. Technology-driven organizations want to capture process and analyze this data at a fast velocity for efficient decision-making to support their operations and their business processes [2, 5-7]. As much as data volume and variety increases and as faster analytic results are needed, more demanding is for data architecture. This data architecture should enable collecting, storing, and analyzing big data [3]. Many companies such as Amazon, Google and Microsoft accelerate their paces in developing cloud-computing systems and enhancing their services to cater for a wide variety of users [4].

The bulk of existing academic literature on social media has been published in just the last few years and has focused on the social processes of social media and its effects in areas such as marketing, politics, health communication, and education [25]. Social media platform such as Twitter has stormed onto the social media scene not only as an individual communication device but also as an information dissemination platform [26]. People on social media express opinions on different topics [27]. There are few studies performing external data competitive analysis on the leading companies in an industry in a systemic way [28]. External data can help decision makers to ensure efficient solutions to the problems raised [27]. However, the trustworthiness of this data is often questionable due to the huge amount of data created. For example, social media has now become an important medium of communication and interaction tools for social networks [29]. Social media is also important for business platform that can influence the corporative environment [30-32]. Social networks involve agents in creating and processing information for knowledge network [33]. At the same time, the role that causality can play in social network analysis is unclear [34]. Therefore, it is important to examine the flow of information share on social media and to retrieve relevant information from large amount of it. While most people take such technologies for granted, our understanding about external data such as social media, mobile data and sensor data are very limited [35]. Most existing studies on this topic remain descriptive, focusing on what people do with it [35, 36]. Despite the abundant research on IT adoption in general, our understanding about the effectiveness of analyzing external data is still at the early stage. Moreover, the uniqueness of this data from the other IT applications may require for information entrepreneur to further their theoretical extension.

A. Cloud for Information Enterprise

Progress of research efforts in a novel technology is contingent on having a rigorous organization of its knowledge domain and a comprehensive understanding of all the relevant components of cloud technology and their relationships [37-40]. Study by Youseff et al. [37] showed that cloud computing is one contemporary technology in which the research community has recently embarked. The technology would enable the organization to design more efficient portals and gateways for analyzing their data and facilitate the adoption of this novel computing approach in scientific environments. For example, cloud storage architecture based on ontology can store and retrieve files in the cloud based on its content [38].

Cloud computing requires scalable and cooperative sharing the resources in various organizations by dynamic configuring a virtual organization according to user's requirements. Study by Yoo et al. [39] showed that ontology-based representation of cloud computing environment would be able to conceptualize common attributes among cloud resources.

Retrieving data from a well organized database is requested to be familiar with its schema, structure, and architecture. An effective approach to retrieve desired information or to extract useful data is an important issue in the emerging cloud [40]. For example, in the mobile cloud environment, the information retrieval system based on the P2P is very important. In case of an information search that is not consistent with the meta profile in the data distribution technique, the reliability of the searched data cannot be assured [41]. According to Surachai & Banditwattanawong [42], there is usually several cloud computing platforms with different features being deployed in the same or different organizations in a multi-cloud environment. This leads to the need to interoperate between different cloud platforms to capture data from large databases.

Table 1. Summary of the previous studies on cloud.

Authors	Onto-logy	Information retrieval	Information Enterprise	External data
Chang et al. [40]	No	Yes	No	No
Golovchinsky et al. [43]	No	Yes	Yes	No
Surachai & Banditwattanawong[42]	Yes	No	Yes	No
McFedries [44]	No	Yes	No	No
Yoo et al. [39]	Yes	No	No	No
Chow et al. [45]	No	Yes	Yes	No

Cloud concept is still changing as a large pool of easily usable and accessible virtualized resources. These resources can be dynamically for an optimum resource utilization [46]. For example, McFedries [44] described data center as the basic unit of cloud. It offers huge amounts of computing power and storage by using spare resources, which relate to the concept of massive data scalability [47]. Most of ongoing works are aiming at developing the techniques and constructing cloud platforms, such as Amazon, Google AppEngine and Microsoft Azure [48]. Massive data in data center of cloud platform can provide benefits to the cloud provider and consumer in retrieving information among business, medical information, and cooperative information retrieval platform [43, 45]. Table 1 shows limited study on cloud that focus on retrieving information from external data. In contrast to our study, we develop ontology on cloud for information retrieval. Based on our observation only Surachai & Banditwattanawong [42] and Yoo et al. [39] incorporate ontology and cloud.

B. *Ontology for Information Enterprise*

Most of the recent studies on ontology focus on system ontology [49-51] and enterprise ontology [52-57]. Some studies include both system ontology and enterprise ontology [58-60] while some studies focus on organizational ontology. In Rao et al. [61], the authors discussed ontology as representing organizational knowledge which provides the means by which to understand the relationships between the organizational goals, sub-goals, business processes, tasks, sub-tasks, resources and decision makers. In this study, the development of ontology is based on knowledge mapping within the enterprise modelling and the author identifies the flow of knowledge within the organization.

On the other hand, Sharma & Osei-Bryson [62] presented an organization-ontology-based framework that not only incorporates the applicable tools and techniques, it also provides the ability to present the output of activities in a form that allows for at least their semi-automated integration with activities of this phase and succeeding phases. However, the authors look at the data mining methodology that is associated with business understanding (BU). This study was followed up by Rao et al. [63], who proposed an approach by building an ontology and a set of corresponding competency questions for the information technology (IT) infrastructure domain. The authors also emphasized that a formal set of ontologies must have a set of formal axioms that provide the basis of the ontology's deductive capability. Mansingh et al. [64] proposed an ontology-driven methodology for extracting different knowledge items and representing them as knowledge maps.

Kang et al. [54] deliberated on enterprise architectures (EA) as an approach to address the problems between humans or between systems or between humans and systems. In order to solve this problem, the authors developed an ontology based on enterprise architecture. Kang et al. [55] also developed a business enterprise ontology and identified the lack of semantics which causes communication problems between humans or between systems or between humans and systems. The authors used a fact-based ontology as a conceptual modelling method to cope with the dynamically changing business environment.

These days, the activities of enterprises are continuously globalized and the business environment is changing rapidly and becoming more complex [52]. In response to the changing business environment, it is important to develop new business models and business processes. Kang et al. [54] developed an ontology based on enterprise modelling to maintain competitiveness by accommodating changes in the business environment quickly and flexibly. However, Kang, Lee & Choi [54] and Kang, Lee & Kim [55] focused on system development based on ontology. Even though these studies are based on an ontology model, the authors do not emphasize any relationship between the internal data and external data in their model.

Another example of enterprise ontology was proposed by O'Leary [57], who provided an alternative, theory-based approach for generating an enterprise ontology using activity theory. The author compares previously proposed enterprise ontologies such as Architecture of Integrated Information Systems (ARIS), Resources, Events, Agents (REA), Toronto Virtual Enterprise (TOVE) and Enterprise Ontology. The author suggests that rather than there being a single centralized organizational ontology, there are multiple partial organizational ontologies scattered among firms.

Jimeno-Yepes et al. [50] studied ontology refinement to improve information retrieval. The authors studied the ontology and terminological resources which have appeared in information retrieval (IR) either to provide query expansion terms, to perform semantic indexing of documents or to assist in the better organization of retrieved documents. However, this ontology is usually not optimal for IR tasks. Han & Park [58] studied an enterprise ontology where knowledge is treated as a critical driving force for attaining the organization's performance goals which is important for decision-making. They proposed a knowledge framework model and an enterprise ontology for the process-centered organizational structure. Huang & Dao [59] focused on the usage of ontology to integrate the knowledge between different domains to improve business processes. The authors believe that organizations are becoming more knowledge intensive, and the integration of various types of knowledge becomes a challenge. Based on this study, an ontology-based workflow is developed to accumulate knowledge in on-going processes and can provide help in complex workflow systems and optimization. Table 2 shows that there is no study focus on cloud that integrates ontology to retrieve information. In this study, we propose ontology on cloud to store and retrieve information from external sources such as social media and mobile data.

Table 2. Summary of the previous studies on ontology.

C. *Information Retrieval for Enterprise*

Authors	Cloud	Information retrieval	Information Enterprise	External data
Rao et al. [61],	No	No	No	No
Huang & Dao [59]	No	No	Yes	No
Jimeno-Yepes et al. [50]	No	Yes	No	No
Sharma & Osei-Bryson [62]	No	No	No	No
Mansingh et al. [64]	No	No	No	No
Han & Park [58]	No	No	Yes	No
O'Leary [57]	No	Yes	Yes	No

There are many information retrieval projects have been proposed [65-70]. Most of the studies show that information retrieval exploiting social media in big data era [29, 33, 35]. Retrieving information is requested to be familiar with its schema, structure, and architecture of social media. However, it is against the concept and characteristics of cloud platforms [40].

In addition, ontology is one of most widely applied approach in enabling system intelligence and improving the automation ability of system by obtaining system semantics, which allow retrieving information can be applied [71-75].

Even though there are studies focuses on cloud and ontology development for managing organizational data and information [33, 39, 40, 42, 43], there is still lack of evidence focuses on the development of ontology to retrieve organizational internal data and external data for efficient decision-making. Table 3 shows that most of the study focuses on cloud and ontology. However, there is still no study that annotates the retrieving process between internal data and external data. In contrast to our research, we integrate cloud and ontology to retrieve data for relevant information.

Table. 3 Summary of the previous studies on information retrieval.

Authors	Cloud	Ontology	Information Enterprise	External data (e.g. social media, mobile data etc)
Chang et al. [40]	Yes	No	No	No
Golovchinsky et al. [43]	Yes	No	Yes	No
Surachi & Banditwattanawong [42]	Yes	Yes	Yes	No
Roth & Cointet [33]	No	Yes	Yes	No
Martino [74]	No	Yes	No	No

In contrast to our study, the main complexity of establishing the proposed mentioned framework lies on the heterogeneity of the data and information sources. In addition, they all have different nature in terms of volume and velocity (e.g. sensor data and social media data may be large amount of streaming data, other data sets may be less dynamic), spatial relativity aspect (e.g. mobile data may have inherent spatial/GPS knowledge that can be extracted), and quality aspects (e.g. crowd-source data may be less reliable), etc.

Analyzing large more complex problem entails using large databases. The framework in this area is designed to make these databases easy to access and to enable domain experts to work easily with the databases. For example, assisting domain experts to retrieve data in a variety of ways or to display data graphically. In order to solve the problems mentioned above, we propose an ontology-based cloud framework to dynamically and fully automatically generate external data for retrieving desired information on cloud platforms. The framework can assist organizations to automatically generate external data for retrieving desired information spread over cloud platforms to fulfill the requirement of domain experts.

III. Proposed Framework

The proposed framework aims to resolve the issue in identifying and evaluating relevant data for better decision-making that covers the characteristics of good quality relevant data. In this research, we propose an Enterprise Global Ontology that can capture the above concepts of complexity and use as a basis for data integration processes suitable for efficient correlative decision-making. This Enterprise Global Ontology will reside on the cloud and it will initially contain the local organizational ontology. We will develop an incremental approach to extend the ontology with relevant external knowledge bases by clustering and annotating the external data to concepts and properties within the ontology.

This organization Global Ontology will reside on the cloud and initially contain the local organizational ontology. The on-premise storage services will house the organizational data and the different off-premise storage services will house the external data which we collectively consider the big data. Our first task is to develop the local organizational ontology to show the flow of data in the organization. After that, we will focus to develop the relationship between local ontology and external data. This local ontology will be extended by clustering the external data to define the degree of relevance of the particular data. The organization Global Ontology will extract knowledge through correlation between local ontology and external data.

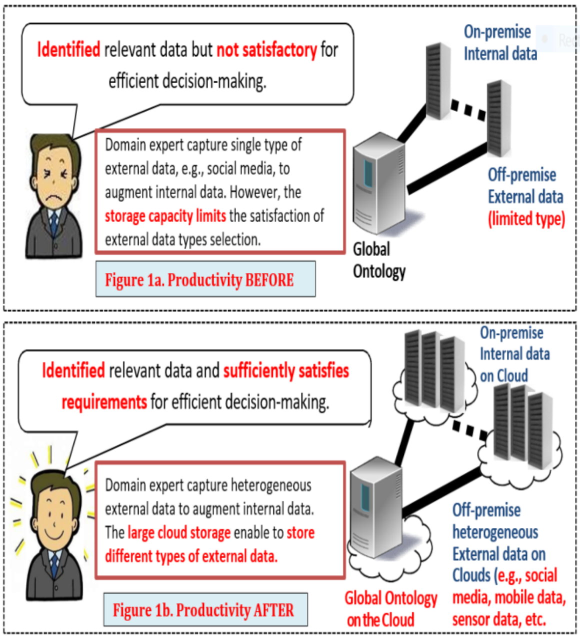

Figure 1. Scenarios with internal data and external data.

The originality of this research lies in the creation of the Big Data integration 'road-map' in the form of an Enterprise Global Ontology that contains references to concepts and properties from external sources relevant to the organization. This road-map will enable accurate integration of the datasets for decision-making as it will serve as a meta-data for the connectivity of the integration process. This approach will enable the organization to have an overall view of data connectivity within and outside the organizations, and to enable the data scientists to harvest interconnected information for analytics purposes.

A. *Application of the Ontology*

The ideas of using an ontology and visual structuring in organization applications were discussed in many works and now are implemented in many sectors [61, 64, 76-78]. However, much of the research in this field did not receive much attention in the literature that incorporate external data such as social media, mobile data and organizational data for efficient decision-making. An ontology provides explicit and formal specifications of knowledge, especially implicit or hidden knowledge [79]. An ontology make the process to identify the relevance of data more easily consumable to address which data from the datasets are more important. The outcome of this research can establish an analytics of Big Data structure for the organizations to ensure that analytics processes are supported by the specific organizational priorities. The contribution of an ontology is to improve the creation of model ultimately takes place through different domain.

In Fig. 2, the Enterprise Global Ontology is built on a cloud computing environment. The on-premise storage services will house the organizational data and the different off-premise storage services will house the external data, which we consider the Big Data. This Big Data is basically dynamic and its relation/connectivity to the organizational data needs to be incrementally and timely updated based on-demand from the user. Enterprise Global Ontology will contain the local organizational ontology and it will be extended by clustering and annotating the relevant external data in a timely manner. The first focus of this research will be on the development of the local organizational ontology to show the flow of data in the organization. After that, we will focus on the relationship between local ontology and external data. This local ontology will be extended by clustering the external data to define the degree of relevance of the particular data. In this research, Enterprise Global Ontology will extract knowledge through the correlation between local ontology and external data. In doing correlation, there are few issues that we want to resolve. For example, finding the most efficient clustering methodology to extract external data. Likewise, it is important to find the best way to link the local organizational ontology to the matched data, which was derived after clustering the external data.

Despite the various initiatives model for the evaluation of organizational process based on ontology [61-64, 80, 81], this paper focus on the ontology to evaluate organizational data that relate to the organizational goals [1, 82]. This process consists of identify whether data is relevant correspond in achieving the organizational goals. This identification focused on domain experts and entrepreneurs who contribute in the decision-making process and responsible to identify to what extend the organizational goals are achieved. The evaluation of ontology shows the dependency relationship of the organizational goals and the dependency relationship of organizational data that relate to the organizational goals. In this section, we propose the Organizational Global Ontology as a model to develop a common understanding between internal data and external data. It provides the domain experts and entrepreneurs with knowledge to identify the most relevant organizational data in relation to organizational goals. In order to propose the organizational goals ontology, we combined several structures that were proposed in previous models [61, 62, 80], adopting these models as a reference during ontology development. However, the scope of our organizational goals ontology does not cover all the organizational processes such as activities, physical resources and performance.

Fox et al. [80] focused on structuring the link between organizational structure and behaviour. This is critical for enterprise model development. However, the authors do not emphasise any organizational resources such as data and information because they focus on the roles and activities within the organization. Meanwhile, Sharma & Osei-Bryson [62] developed a framework for an organizational ontology in an effort to increase an understanding of the business. However, in this study, the authors do not specifically identify the relationship between organizational resources, such as data, and the organizational goals. In this model, the authors adapted the work of Fox et al. [80], where the authors discussed the physical resources and role of the organizational model. Recently, Rao et al. [61] developed an organizational ontology in order to build a knowledge map within the organization. The structure includes the flow of knowledge within the organization in the context of knowledge sharing and knowledge storage. In this model, the authors discussed the organizational resources, as in Sharma & Osei-Bryson [62]. Another aspect that is similar to Sharma & Osei-Bryson's work is that both models include business processes. However, Rao et al. [61] discussed business processes from the organizational goal point of view and Sharma & Osei-Bryson [62] discussed business processes from the organizational activity point of view. Most previous studies focused on organizational structure and performance. To our knowledge, very few studies have been conducted on the development of the organizational goals ontology in an effort to evaluate the relationship between organizational data and organizational goals.

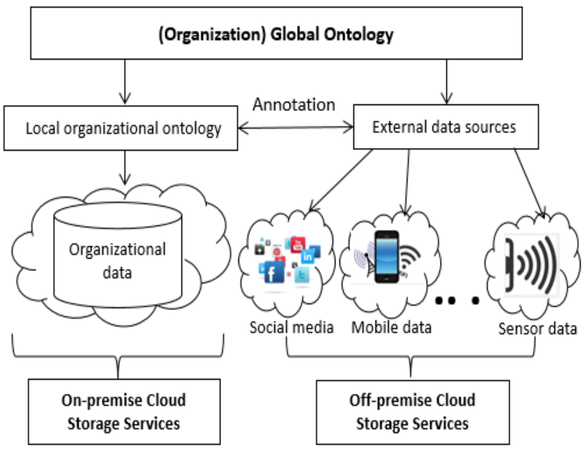

Figure 2. Proposed semantic-based Big Data correlation on cloud.

B. *Organization Global Ontology*

Ontology architecture as presented in Fig. 3 consists of ontology design and ontology application. Ontology design is a stage where we apply the ontology to identify the goals and different types of data from external data and internal data sources. After we identify the data, we develop the relationship between data and organizational goals.

Ontology application is a stage where we analyse this data to support decision-making process in relation to the organizational goals. We come out with the decision-making process to evaluate the organizational goals achievement. In this paper, an ontology is applied for big data to:

- be applicable in a wide range of domain.
- successfully develop the dependency relationship between different types of data.
- successfully develop the dependency relationship between data and goals.

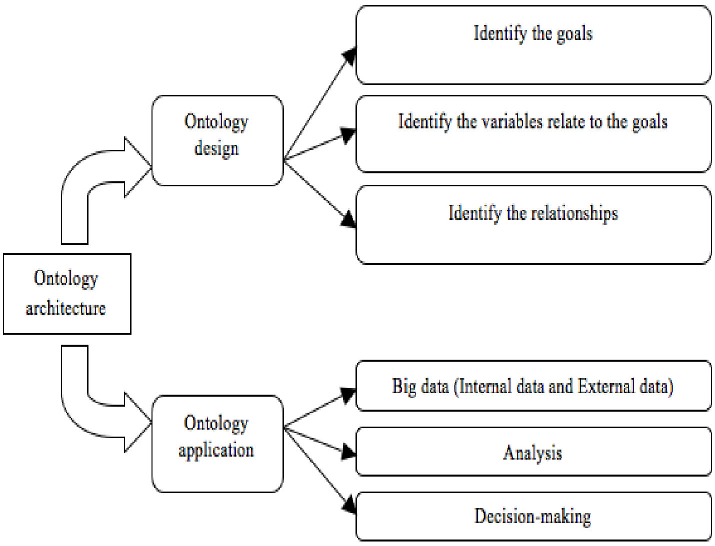

Figure 3. Ontology architecture.

In this paper, Organization Global Ontology focuses on the organizational goals. Organizational goals are defined as the most important targets to be achieved in every organization [1]. Even though the concept of the organizational goals has been in the existence for some time, modeling the structure of the organizational goals is much more difficult [1, 82]. For example, one way to develop a common understanding of the organizational goals structure is based on an ontology [1]. An ontology provides explicit and formal specifications of knowledge, especially implicit or hidden knowledge [79]. An ontology is considered as an approach to support data sharing [83]. Therefore, an ontology assists with part of the integration problem in relation to the organizational goals and can be used to improve the communication and collaboration between the decision makers and the users [84].

In Izhar et al [1], organizational goals ontology is developed based on the work of Rao et al. [61], Sharma & Osei-Bryson [62] and Fox et al. [80]. Despite many research efforts and established model for the organizational goals using an ontology, they have not yet been systematically applied for decision-making to support the evaluation of the organizational goals achievement. This is important because decision-support is one of the main objectives of an ontology [85]. In this paper, we extend the organizational goals ontology developed in Izhar et al. [1], in order to develop the relationship between external data and internal data.

Several structures that were proposed in the previous models are combined [61, 62, 80] for the organizational goals ontology in Izhar et al. [1]. These models are adapted as a reference for the organizational goals ontology. However, the scope of the proposed organizational goals ontology in this methodology do not cover all the organizational processes as discussed in Sharma & Osei-Bryson [62], Fox et al. [80] and Rao et al. [61].

Fox et al. [80] focused on structuring the linkage between organizational structure and behavior. This is critical for enterprise model development. However, the authors do not emphasize any organizational resources such as data and information but they focus on the roles and activities within the organization. Meanwhile, Sharma & Osei-Bryson [62] developed a framework for an organizational ontology in an effort to increase an understanding of the business. However, the authors do not specifically identify the relationship between organizational resources, such as data and the organizational goals. In this model, the authors adapted the work of Fox et al. [80], where the authors discussed the physical resources and role of the organizational model.

Recently, Rao et al. [61] developed an organizational ontology in order to build a knowledge map within the organization. The structure includes the flow of knowledge within the organization in the context of knowledge sharing and knowledge storage. In this model, the authors discussed the organizational resources, as in Sharma & Osei-Bryson [62]. Another aspect that is similar to Sharma & Osei-Bryson's work is that both models include business processes. However, Rao et al. [61] discussed business processes from the organizational goals point of view and Sharma & Osei-Bryson [62] discussed business processes from the organizational activity point of view. Most of these studies focused on the organizational structure and performance. In Izhar et al. [1], the authors developed the organizational goals ontology that consists of organizational goals, sub-goals, and organizational data. They developed the dependency relationship for the organizational goals and dependency relationship between organizational data and organizational goals. However, they evaluate the organizational goals by identifying the organizational goals first and then they identify the organizational data that relate to the organizational goals.

In organizations, it is extremely important for the manager to have access to the most relevant data in relation to the organizational goals [1]. Simsek et al. [86] 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 domain experts and entrepreneurs to use the most relevant data to successfully achieve the organizational goals.

Table 4. Scope on organizational goals ontology forinternal data and external data

Authors	Organizational goals ontology		Resources	
	Organizational goals	Sub-goals	Internal Data	External Data
Fox el al. [80]	✓	✓	×	×
Sharma & Osei-Bryson [62]	✓	✓	×	×
Rao et al. [61]	✓	✓	×	×

Izhar et al. [1]	✓	✓	✓	×

IV. Expected outcomes

The expected contribution of the framework will benefit big data challenges for accurate data analysis. Big data may be as important to business because more data can lead to more accurate analyses. More accurate analyses may lead to more confident decision-making and better decision can mean greater operational efficiencies in the organization [8].

Big data is a new way of thinking about enterprise data and how it can drive business value. The amount of data that is available to businesses is increasing, with social media and machine-to-machine as just two of the leading sources. The central role of business services in today's enterprises, and the more complex architecture through which they are delivered, make it important to manage big data solutions from a business perspective. Business perspective focuses on business objectives and benefit, and prioritizes resources and activities according to the needs of the business. In this way, effective evaluation of the big data can ensure optimal relevance of data for more effective decision-making to support the business goals.

Taking advantage of big data opportunities is challenging for the organizations [2]. Firms and other organizations have been using large databases and analytics for the last couple of decades. Transactions are stored in data warehouses and analyzed with data-mining algorithms to extract insights [5]. In order to ensure the effectiveness of the data, organizations need to be able to store data reliably across a number of databases. Once data need to be distributed, organizations need a way to get it out again and they need to identify which data is needed, assemble it and analyse it. The challenge is how to capture this data to be considered relevant for the specific organization activities because determining relevant data is a key to delivering value from massive amounts of data as shown in Fig. 2. The real issue is not how the organizations acquiring large amount of data but how they do with the data that counts [8]. The technologies and concepts behind big data can allow organizations to achieve a variety of objectives.

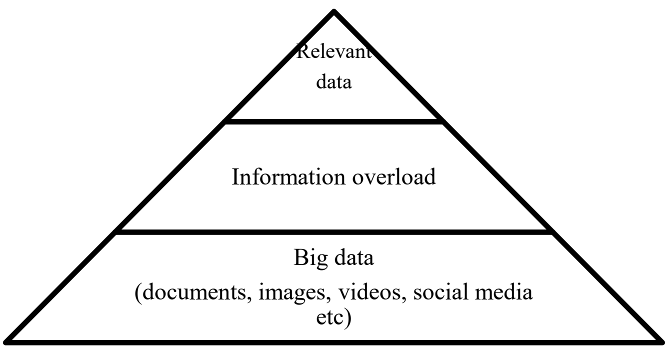

Figure 4. Relevant data from large amounts of data.

Impact on Society – Many workplaces in general creates huge amount of data. This study will improve the productivity in terms of knowledge management implication when dealing with large volume of data. Such workplaces will not only practice values of managing data, information and knowledge but also to promote trustworthiness of data and information in multi-disciplinary background that advocates these values to the society.

Impact on the Economy – Having an ability to manage data is crucial ingredient for k-economy that helps to transform a part of human knowledge to machines by using particular knowledge technology. This technology will improve performance in dealing with data, work performance and efficient decision-making delivery for more competitive and productive economy. Therefore, having the ability to analyze data in a timely fashion can ensure the ability to improve productivity in relation to decision-making. This research is multi-disciplinary that could find correlation between data in different domains such as disaster management and people movements, biology and agriculture sciences, health care and population, etc.

Impact on the Nation - The results can lead to some new evaluation methods in big data era from different perspectives. The outcome will offer an enormous opportunity to advance the science of data analytics so that future researchers will have a new understanding on what is needed to improve their data analysis process. The conduct of this research will ensure the projects agility in responding to unfolding events, and substantially enhance its ability to engage in and impact on societies and nation.

V. Conclusion

The ability to draw correlation and extract knowledge from relevant internal and external data will provide accurate situational awareness that supports effective decision-making. The outcome of this research includes (i) consolidated decision-making process, (ii) cloud services for data matching and integration techniques, and (iii) incremental global ontology expansion. Beside business enterprise, the Global Ontology concept can also be applied to other areas which are highly impacted by things that are external to the organization (environmental conditions, weather, public sentiments, etc.), such as the agricultural industry, health care systems, transport industry, animal husbandry, and other industries in which decision-making is critical.

This research will lead to some experimental results that can lead to new methods and innovation in analysing data from different perspectives. The outcome will offer an enormous opportunity to advance the science of data analysis so future researchers have new understanding of what is needed to improve their data analysis process. The presence of this research ensure the projects agility in responding to unfolding events, and substantially enhance its ability to engage in and impact on the development of better analytic tool to manage large and complex databases.

References

[1] T.A.T. Izhar, T. Torabi, M.I. Bhatti, F. Liu, Recent developments in the organization goals conformance using ontology, Expert Systems with Applications, 40 (2013) 4252-4267.
[2] M. Berber, E. Graupner, A. Maedche, The information panopticon in the big data era, Journal of Organization Design, 3 (2014) 14-19.
[3] J.O.e. Sá, C. Martins, P. Simões, Big Data in Cloud: A Data Architecture, New Contributions in Information Systems and Technologies, 353 (2015) 723-732.
[4] J. Kang, K.M. Sim, Ontology and search engine for cloud computing system, IEEE 2011 International Conference on System Science and Engineering (ICSSE), 2011, pp. 276-281.
[5] J.R. Galbraith, Organization design challengers resulting from big data, Journal of Organization Design, 3 (2014) 2-13.
[6] B.T. Hazen, C.A. Boone, J.D. Ezell, L.A. Jones-Farmer, Data quality for data science, predictive analytics and big data in supply chain management: An introduction to the problem and suggestions for research and application, International Journal Production Economics, 154 (2014) 72-80.
[7] R.L. Grossman, K.P. Siegel, Organizational models for big data and analytics, Journal of Organization Design, 3 (2014) 20-25.
[8] T.H. Davenport, J. Dyche, Big data in big companies, International Institute for Analytics, 2013.
[9] J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, A.H. Byers, Big data: The next frontier for innovation, competition and productivity, McKinsey Global Institute 2011.
[10] S. Lohr, How big data became so big, New York Times, 2012.
[11] H. Chen, R.H.L. Chiang, V.C. Storey, Business intelligence and analytics: from big data to big impact, MIS Quarterly, 36 (2012) 1165-1188.
[12] R.K. Perrons, J.W. Jensen, Data as an asset: Why the oil and gas sector can learn from other industries about "Big Data", Energy Policy, 81 (2015) 117-121.
[13] Q. Zhi, L. Zhao-Wen, M. Yan, Research of Hadoop-based data flow management system, The Journal of China Universities of Post and Telecommunications, 18 (2011) 164-168.
[14] C. Esposito, M. Ficco, F. Palmieri, A. Castiglione, A knowledge-based platform for big data analytics based on publish/subscribe services and stream processing, Knowledge-Based Systems, 79 (2015) 3-17.
[15] D.G. Omerzel, B. Antoncic, Critical entrepreneurs knowledge dimensions for the SME performance, Industrial Management and Data System, 8 (2008) 1182-1199.
[16] K. Schalenkamp, W.L. Smith, Entrepreneurial skills assessment: the perspective of SBDC directors., International Journal of Management and Entreprise Development, 5 (2008) 18-29.
[17] W.L. Smith, K. Schalenkamp, D.E. Eicholz, Entrepreneurial skills assessment: an exploratory study, International Journal of Management, 4 (2007) 179-201.
[18] A. Mikroyannidis, B. Theodoulidis, Ontology management and evolution for businee intelligence, International Journal of Information Management, 30 (2010) 559-566.
[19] D. Xiaoying, L. Qianqian, Y. Dezhi, Business performance, business strategy, and information system strategic alignment: An empirical study on Chinese firms, Tsinghua Science and Technology, 13 (2008) 348-354.
[20] P. Trkman, The critical success factors of business process management, International Journal of Information Management, 30 (2010) 125-134.
[21] S. Jonsson, J. Lindbergh, The impact of institutional impediments and information and knowledge exchange on SMEs' investments in international business relationships, International Business Review, 19 (2010) 548-561.
[22] N. Harmancioglu, A. Grinstein, A. Goldman, Innovation and performance outcomes of market information collection efforts: The role of top management team involvement, International Journal of Research in Marketing, 27 (2010) 33-43.
[23] M.Z. Elbashir, P.A. Collier, M.J. Davern, Measuring the effects of business intelligence systems: The relationship between business process and organizational performance, International Journal of Accounting Information Systems, 9 (2008) 135-153.
[24] J. Barjis, The importance of business process modeling in software systems design, Science of Computer Programming, 71 (2008) 73-87.
[25] K. McIntyre, The evolution of social media from 1969 to 2013: A change in competition and a trend toward complementary, The Journal of Social Media in Society, 3 (2014) 6-25.
[26] D. Krause, M. Smith, Twitter as mythmaker in storytelling: the emergence of hero status by the Boston police department in the aftermath of the 2013 marathon bombing, The Journal of Social Media in Society, 3 (2014) 8-27.
[27] N.F.F.d. Silva, E.R. Hruschka, E.R.H. Jr, Tweet sentiment analysis with classifier ensembles, Decision Support Systems, 66 (2014) 170-179.
[28] W. He, S. Zha, L. Li, Social media competitive analysis and text mining: A case study in pizza industry, International Journal of Information Management 33 (2013) 464-472.
[29] N.A. Ghani, S.S.M. Kamal, A Sentiment-based filteration and data analysis framework for social media, 5th International Conference on Computing and Informatics ICOCIIstanbul, Turkey, 2015, pp. 632-637.
[30] L. Holtzblatt, J.L. Drury, D. Weiss, L.E. Damianos, D. Cuomo, Evaluating the uses and benefits of an enterprise social media platform, Journal of Social Media for Organizations, 1 (2013) 1-21.
[31] L.E. Domianos, D. Cuomo, J. Griffith, D.M. Hirst, J. Smallwood, Exploring the adoption, utility and social influences of social bookmarking in a corporate environment, International Conference on System SciencesHawaii, 2007.
[32] L.E. Damianos, D.L. Cuomo, S. Drozdetski, Handshake: A case study for exploring business networking for enterprise, inside and out, International Conference on Human Computer Interaction, 2011.
[33] C. Roth, J.P. Cointet, Social and semantic coevolution in knowledge networks, Social Networks, 32 (2010) 16-29.
[34] P. Doreian, Causality in social network analysis, SOCIOLOGICAL METHODS & RESEARCH, 30 (2001) 81-114.
[35] T.A. Pempek, Y.A. Yermolayeva, S.L. Calvert, College students' social networking experiences on Facebook, Journal of Applied Developmental Psychology, 30 (2009) 227–238.
[36] K. Subrahmanyama, S.M. Reich, N. Waechter, G. Espinoza, Online and offline social networks: Use of social networking sites by emerging adults, Journal of Applied Developmental Psychology, 29 (2008) 420–433.
[37] L. Youseff, M. Butrico, D.D. Silva, Toward a unified ontology of cloud computing, IEEE Grid Computing Environments Workshop, (2008) 1-10.
[38] H.T. Al-Feel, M.H. Khafagy, Ocss: Ontology cloud storage system, 2011 First International Symposium Network Cloud Computing and Applications (NCCA) 2011, pp. 9-13.
[39] H. Yoo, C. Hur, S. Kim, Y. Kim, An ontology-based resource selection service on science cloud, Springer Berlin Heidelberg2009.
[40] Y.-S. Chang, C.-T. Yang, Y.-C. Luo, An Ontology based Agent Generation for Information Retrieval on Cloud Environment, Journal of Universal Computer Science, 17 (2011) 1135-1160
[41] S.-J. Moon, C.-P. Yoon, Information retrieval system using the keyword concept net of the P2P service-based in the mobile cloud environment, Peer-to-Peer Networking and Applications 8(2015) 596-609.
[42] H. Surachai, T. Banditwattanawong, An Interoperability Ontology for Multi-Cloud Computing Platforms, The 10th International Conference on e-Business, 2015, pp. 1-4.
[43] G. Golovchinsky, P. Qvarfordt, J. Pickens, Collaborative Information Seeking, IEEE Computer 42 (2009) 47-51.
[44] P. McFedries, The cloud is the computer. IEEE Spectrum Online, August 2008, 2008.
[45] T. Chow, T. W.-T, B. J, X. Bai, Ontology-based Information Sharing in Service-Oriented Database Systems, IEEE International Conference on Services Computing, 2009, pp. 276-283.
[46] L.M. Vaquero, L. Rodero-Merino, J. Caceres, M. Lindner, A Break in the Clouds: Towards a Cloud Definition, SIGCOMM Computer Communication, 39 (2009) 50-55.
[47] E. Hand, Head in the clouds, Nature, 2007, pp. 963.
[48] C. Vecchiola, X. Chu, R. Buyya, Aneka: A Software Platform for .NET-based Cloud Computing, 2009.
[49] G. Beydoun, A.A. Lopez-Lorca, F. Garcia-Sanchez, R. Martinez-Bejar, How do we measure and improve the quality of a hierarchical ontology?, The Journal of System and Software, 84 (2011) 2363-2373.
[50] A. Jimeno-Yepes, R. Berlanga-Llavori, D. Rebholz-Schuhmann, Ontology refinement for improved information retrieval, Information Processing and Management, 46 (2010) 426-435.
[51] C. Chandra, A. Tumanyan, Organization and problem ontology for supply chain information support system, Data & Knowledge Engineering, 61 (2007) 263-280.
[52] J.-H. Park, K.-H. Kim, J.-H.J. Bae, Analysis of shipbuilding fabrication process with enterprise ontology, Computers in Human Behavior, 27 (2011) 1519-1526.
[53] B. Sharp, A.S. Atkins, H. Kothari, An ontology based multi-agent system to support HABIO outsourcing framework, Expert Systems with Applications, 38 (2011) 6949-6956.
[54] D. Kang, J. Lee, S. Choi, K. Kim, An ontology-based enterprise architecture, Expert Systems with Applications, 37 (2010) 1456-1464.
[55] D. Kang, J. Lee, K. Kim, Alignment of business enterprise architectures using fact-based ontologies, Expert Systems with Applications, 37 (2010) 3271-3283.
[56] M. Zacarias, H.S. Pinto, R. Magalhaes, J. Tribolet, A 'context-aware' and agent-centric perspective for the alignment between individuals and organizations, Information Systems, 35 (2010) 441-466.
[57] D.E. O'Leary, Enterprise ontologies: Review and an activity theory approach, International Journal of Accounting Information Systems, 11 (2010) 336-352.

[58] K.H. Han, J.W. Park, Process-centered knowledge model and enterprise ontology for the development of knowledge management system, Expert Systems with Applications, 36 (2009) 7441-7447.
[59] N. Huang, S. Diao, Ontology-base enterprise knowledge integration, Robotics and Computer-Integrated Manufacturing, 24 (2008) 562-571.
[60] V. Anaya, G. Berio, M. Harzallah, P. Heymans, R. Matulevicius, A.L. Opdahl, H. Panetto, M.J. Verdecho, The Unified Enterprise Modelling Language-Overview and further work, Computers in Industry, 61 (2010) 99-111.
[61] L. Rao, G. Mansingh, K.-M. Osei-Bryson, Building ontology based knowledge maps to assist business process re-engineering, Decision Support Systems, 52 (2012) 577-589.
[62] S. Sharma, K.-M. Osei-Bryson, Organization-ontology based framework for implementing the business understanding phase of data mining projects, International Conference on System SciencesHawaii, 2008, pp. 27.
[63] L. Rao, H. Reichgelt, K.-M. Osei-Bryson, An approach for ontology development and assessment using a quality framework, Knowledge Management Research and Practice, 7 (2009) 260-276.
[64] G. Mansingh, K.-M. Osei-Bryson, H. Reichgelt, Building ontology-based knowledge maps to assist knowledge process outsourcing decisions, Knowledge Management Research and Practice, 7 (2009) 37-51.
[65] Cannataro, D. Talia, T. Paolo, Distributed data mining on the grid, Future Generation Computer Systems, 18 (2002) 1101-1112.
[66] Cannataro, A. Congiusta, A. Pugliese, D. Talia, P. Trunfio, Distributed data mining on grids: Services, tools, and applications, IEEE Transactions on Systems, Man, and Cybernetics, 34 (2014) 2451-2465.
[67] S. Krishnaswamy, S.W. Loke, A. Zaslasvky, A hybrid model for improving response time in distributed data mining, IEEE Transactions on Systems, Man and Cybernetics Part B 34 (2004) 2466 -2479.
[68] V. Stankovski, M. Swain, V. Kravtsov, T. Niessen, D. Wegener, J. Kindermann, W. Dubitzky, Grid-enabling data mining applications with DataMiningGrid: An architectural perspective, Future Generation Computer Systems 24 (2008) 259-279.
[69] G. Wang, T. Wen, Q. Guo, X. Ma, A Knowledge Grid Architecture Based on Mobile Agent, Second International Conference on Semantics, Knowledge, and Grid, 2006, pp. 48 – 51.
[70] W.-C. Shih, C.-T. Yang, S.-S. Tseng, Performance-based Data Distribution for Data Mining Applications on Grid Computing Environments, Journal of Supercomputing 52 (2010) 171-198.
[71] M. Bhatt, A. Flahive, C. Wouters, W. Rahayu, D. Taniar, MOVE: A Distributed Framework for Materialized Ontology View Extraction, Algorithmica, 45 (2006) 457-481.
[72] M. Bhatt, A. Flahive, C. Wouters, W. Rahayu, D. Taniar, T. Dillon, A Distributed Approach to Sub-Ontology Extraction, AINA, 1 (2004) 636-641.
[73] M. Bhatt, C. Wouters, A. Flahive, W. Rahayu, D. Taniar, Semantic Completeness in Sub-ontology Extraction Using Distributed Methods, ICCSA, 3 (2004) 508- 517
[74] B.D. Martino, Semantic web services discovery based on structural ontology matching, International Journal of Web and Grid Services 5(2009) 46-65.
[75] M. Lanzenberger, J. Sampson, M. Rester, Ontology Visualization: Tools and Techniques for Visual Representation of Semi-Structured Meta-Data Journal of Universal Computer Science 16 (2010) 1036-1054.
[76] J. Valaski, A. Malucelli, S. Reinehr, Ontologies application in organizational learning: A literature review, Expert Systems with Applications, 39 (2012) 7555-7561.
[77] M.B. Almeida, R.R. Barbosa, Ontologies in knowledge management support: a case study, Journal of the American Society for Information Science and Technology, 60 (2009) 2032-2047.
[78] M.-C. Valiente, E. Garcia-Barriocanal, M.-A. Sicilia, Applying an ontology approach to IT service management for business-IT integration, Knowledge-Based Systems, 28 (2012) 76-87.
[79] J. Cho, S. Han, H. Kim, Meta-ontology for automated information integration of parts libraries, Computer-Aided Design, 38 (2006) 713-725.
[80] M.S. Fox, M. Barbuceanu, M. Gruninger, J. Lin, An organization ontology for enterprise modelling, Simulation organizations: Computational models of institutions and groupsAAAI/MIT Press1998, pp. 131-152.
[81] M.S. Fox, M. Barbuceanu, M. Gruninger, An organisation ontology for enterprise modeling: Preliminary concepts for linking structure and behaviour, Computers in Industry, 29 (1996) 123-134.
[82] T.A.T. Izhar, T. Torabi, I. Bhatti, F. Liu, Analytical dependency between organisational goals and actions: Modelling concept, International Conference on Innovation and Information Management (ICIIM 2012) IACSIT Press, Chengdu, China, 2012.
[83] H. Pundt, Y. Bishr, Domain ontologies for data sharing-an example from environmental monitoring using field GIS, Computer & Geosciences, 28 (2002) 95-102.
[84] K. Selma, B. Ilyes, B. Ladjel, S. Eric, J. Stephane, B. Michael, Ontology-based structured web data warehouses for sustainable interoperability: requirement modeling, design methodology and tool, Computer in Industry, 63 (2012) 799-812.
[85] A.S. Bastinos, M. Krisper, Multi-criteria decision making in ontologies, Information Sciences, 222 (2013) 593-610.
[86] Z. Simsek, M.H. Lubatkin, J.F. Veiga, R.N. Dino, The role of an entrepreneurially alert information system in promoting corporate entrepreneurship., Journal of Business Research., 62 (2009) 810-817.

Author Biographies

Tengku Adil Tengku Izhar, Ph.D., is a lecturer at the Faculty of Information Management, Universiti Teknologi MARA, Malaysia. His teaching and research interests are in big data, ontology, information management, social media and organizational knowledge assets.

Prof. Bernady O. Apduhan, is Senior Lecturer at the Faculty of Information Science, Kyushu Sangyo University, Japan. His research interests include software engineering, ontology, information systems and computer science.