Institute of Advanced Engineering and Science

International Journal of Information & Network Security (IJINS)
Vol.2, No.5, October 2013, pp. 371~384
ISSN: 2089-3299

On Information Overload Subjugation: Proposing an Intelligent Data Warehouse (IDW) Model

Patrick Mukala, Prof.Van der Walt
Department of Computer Science, Faculty of ICT, Tshwane University of Technology

Article Info

Article history:

Received Jun 12th, 2013
Revised Jul 20th, 2013
Accepted Aug 26th, 2013

Keyword:

Information Overload,
Information Processing, Data Warehousing,
Agent Technology,
Intelligent Data Warehouse (IDW)

ABSTRACT

In this information era, information overload poses a major threat for critical decision making in organizations. From airline industries and manufacturing companies getting the right information at the right time has constituted a major factor for business growth and success. This requires efficient data management tools for information processing. The quest for improved data management tools has been at the heart of data engineering research for many years. However, in spite of such efforts, organizations inevitably and constantly continue to hit the information overload wall. In this paper, we propose a way of using agent technology in data warehouses in order to mitigate the effects of information overload. We propose a layering model called Intelligent Data Warehouse (IDW) for data processing and information retrieval optimization. We give an example of how communication would occur if this model were applied, as well as how ontology for the model could be developed. Finally, we use a case study to simulate aspects of the model, such as the interactions between agents in a context of multi-agent environments, in order to illustrate the occurrence of delegation and information-processing when the suggested model is applied.

Corresponding Author:

First Author,
Department of Computer Science, Faculty of ICT, Tshwane University of Technology
Email: patrick.mukala@gmail.com

1. Introduction

In today's information era, organisations are faced with the task of successfully managing increasing volumes of data [10] and, more importantly, turning these data into meaningful information that is vital to their business operations [6]. These organisations are driven by spectacular technological changes in terms of Information Technology infrastructures and the key challenge remains to deliver secure, high-performance, high-availability access to critical data [14].This requires the use of efficient technologies to cater for effective data manipulation, in order to produce required useful business information. From storage to processing, operations data require effective technologies and tools to produce resulting useful information [11].

Given the speed of current technological advances, information needs are increasing within these organisations and this has made it even more complex to provide users with relevant and up-to-date information. This information results naturally from a rapidly-growing amount of information (mostly stored in data warehouses) which becomes overwhelmingly difficult to analyse.

Several ways used to improve the efficiency (query response time) of data warehouses include appropriate indexing, materialised views, data partitioning and parallel processing. However, these methods have not always been entirely successful in finding desired information within a relatively short period of time, because of massive volumes of unnecessary data [1]. As such, the data warehouse can be perceived, due to this factor, as a constrained environment [12].

Hence, in this paper we consider the possibility to improve data warehouses as data storage repositories using the benefits of agent technology. Tremendous benefits of Agent technology include the ability for information analysis and filtering [2], monitoring processes and operations such as data brokering, assistant agent, process-simulations, etc. [2] The rest of the paper is divided as follows: In section 2, we discuss this approach called Intelligent Data Warehouse (IDW) and give future perspectives in section 3.

2. Intelligent Data Warehouse (IDW) *Model:* quid est?

Enough literature has been published to illustrate the effects of information overload on strategic decision making in organizations. Hence, it is crucial for novel ways of information processing be introduced. Filtering these data and providing the required information as insights for decision-making is very critical and important for organizations.

An Intelligent Data Warehouse (IDW) model is a model that integrates intelligent functionalities such as improved data mining, effective data gathering and processing based on filtering criteria in accordance with the user specifications and requirements, in order to reduce the query response time despite volumes of stored data.The fundamental approach of the model is to integrate special features for the storage of data, the processing or cleaning of this data and, most importantly, the improved retrieval of this information.

The Intelligent Data Warehouse is one that can be used to successfully uncover the meanings about hidden relationships among data. Information overload and time-consuming data accessing in ordinary data warehouses occur due to ambiguity in stored datasets and clearly because of traffic caused during distribution. The ambiguity in datasets is a result of the different formats used in the queries performed for data retrieval. Data warehouses contain large amounts of information retrieved from diverse sources which include databases and relational databases, as well as external sources. This collected data undergoes intensive cleaning or processing in order to produce a specific requested output. When the formats of the data in these databases do not match the queries' format, there is miscommunication and incompatibility of data, thus creating ambiguity.

Distribution, on the other hand, is an issue that existing data warehouses face because of concurrent and simultaneous data requests. When multiple users or entities such as decision-support systems require specific information from a data warehouse at the same time, the data warehouse is overloaded and thus it is unable to produce the relative required outputs in due time.

Data warehouses perform high-level analytical operations that are unknown or not even comprehensible to the end-user. Furthermore, managing these multiple data sources requires the use of a specialised system to coordinate operations across a network.

An Intelligent Data Warehouse is intended to integrate agent-based functionality that enables the treatment of data from these different sources locally and then performs accurate and timely information distribution when requested in a delegation-system-like operating mode.

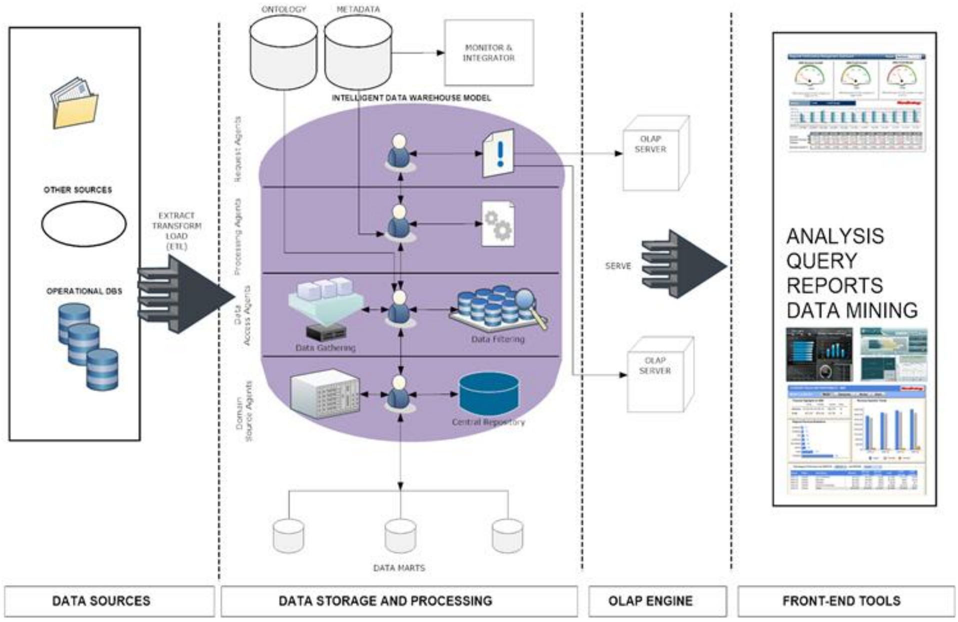

Figure 1. Intelligent Data Warehouse (IDW) model

2.1. Architecture and Functionalities

In order to allow efficient data integration and to cater for high-performance information distribution, the virtual architecture of this approach will encompass all the main components of a typical data warehouse including specific variations at the core of the structure which include four layers. Based on the ontological design and semantics of agent technology, Figure 1 depicts a graphical representation of the proposed model. Notice that the model is an improvement of existing data warehousing design.

The architecture comprises the main data warehouse components such as back-end tools and utilities that extract transform and load data from the data sources (operational databases, etc.) to the data storage and processing component, which is the core of the data warehouse. Here, data can be stored using a relational data model (e.g., star schema) or a multi-dimensional data model implemented by arrays. This is where the originality of the model can be noticed. We have introduced a four-layered structure to extract and load data using agent technology:

a. Request Layer: This is the part of the intelligent data warehouse that deals with the graphical user interface. The GUI of the data warehouse, designed to help the user specify the search parameters, incorporates a set of agents that will help with receiving user-supplied information, format it accordingly and forward it to the next agents in charge of processing the requests and rendering the required services. User request agents are also responsible for providing the user with processed information received from the processing agents.
b. Processing Layer: These are agents that receive formatted requests from the request agents and perform all related operations in order to provide expected outputs. However, these agents need data-related to user-specified requests in order to provide the requested output. They analyse the parameters or requests from the user and then identify all types of data that they will need to render the requested service, thus they request in turn a set of data from data access agents using the metadata base.
c. Data Access Layer: These agents execute the information search operations. The operations involve executing a data search process based on specified parameters from the processing agents. Enough data needs to be gathered to satisfy the user search, thus these agents will conduct data-gathering operations as well as data-filtering operations to ensure that the requested results include insights from a various range of sources making use of the ontology base to reduce knowledge disparity. Furthermore, as this information emanates from multiple distributed environments or data sources, these agents also send requests to the last cluster of agents which base their search on which data comes from where, identifying the diverse integrated sources from which to obtain the requested information.
d. Domain Source Layer: These agents enable the search by specifying the domain of the distributed data sources. Domain-based agents conduct lower-level and technical tasks to integrate the information from these sources and channel them to the central repository for storage.

2.2. IDW's Ontology Definition

Our model, the Intelligent Data Warehouse Model (IDW) proposes an integration environment. And in such environment, the knowledge disparity existing between information sources is critical. Each of these sources belongs to a specific entity and structure with specific standards and functional operations. This can turn the environment's operating process very complex and communication non-existent. Hence, to avoid such a situation, IDW as a model for the implementation of an integration environment will avoid this situation and reduce and/or eliminate this disparity in order to reduce the complexity of operation which involves multiple heterogeneous and distributed entities. The model will allow for handling of the issue of knowledge disparity through Ontology through the set of mapping rules and related details such as meaning, implementation, format etc.

Ontology in the context of knowledge-sharing refers to a specification of a conceptualisation. That is, ontology is a description (like a formal specification of a program) of the concepts and relationships that can exist for an agent or a community of agents. Ontologies are designed for the purpose of enabling knowledge-sharing and re-use. It is a set of definitions of formal vocabulary as described in figure 2 below.

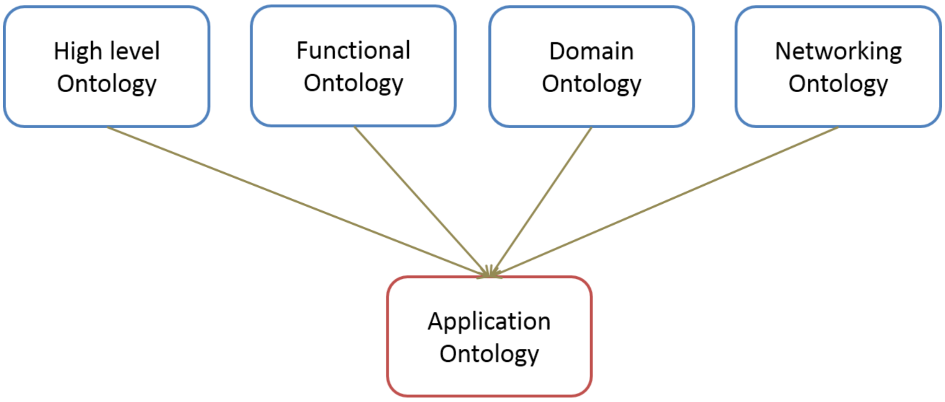

Figure 2. Collection of Ontologies in Intellligent Agent Tourism Environment

For example, a tourist preparing for a journey suffers from information overload when they use the internet to plan their next holiday.

One of the core goals of the Semantic Web is to store data in distributed locations, and use ontology's and reasoning to aggregate it. The current problem is that there exists currently much ontology designed to solve a variety of semantic problems. As shown in figure 2, we have defined a collection of ontologies for modelling the context in an intelligent agent tourism environment; expressed in the Web Ontology Language OWL. The ontology's defines typical concepts associated with the domain (tourism}, function (data warehousing) and networking (collaboration) to support and integrate the main application.

Using the Ontology base, the third layer of the model specifies how the several sources of data are semantically linked and how the integration operations are to be performed in real time. The IDW's ontology base can be defined as incorporating knowledge about four important components. These components include the component/entity structure, mapping rules, semantic links as well as software structure.

On the entity/component structure, the ontology base contains knowledge about the inventory of the component used in the schema of each source, such as the entity names, the relationship/association names, the attribute name, the attribute types, etc.; the mapping rules explain the process used to map an information from one format to another, such as conversion details form a currency to another , or for example from the inch to the centimetre, these rules also incorporate existing data warehouse rules; semantic links specify the structures that are different at the format level, but are the same at the meaning level; and finally, the software structure details the tools used in the implementation of each source, such as the programming language, DBMS etc.

For illustrative purpose, below we use a case study to exemplify how to define and create ontology using OWL. OWL, as previously indicated, is an ontology language that formally describes the meaning of terminology used in Web documents. It extends from and is a revision of the DARPA Agent Markup Language+Ontology Interface Layer (DAML+OIL) language.

2.2.1 Ontology for a Movie event

To design our ontology, we made use of a modelling software application called Protégé and we used the following iterative methodology:

- Step1: Determine the domain and scope of the ontology
- Step 2: Consider re-using existing ontologies
- Step 3: Enumerate important terms in the ontology
- Step 4: Define the classes and the class hierarchy
- Step 5: Define the properties of classes - slots
- Step 6: Define the facets of the slots
- Step 7: Create Instances

STEP 1: DETERMINE THE DOMAIN AND SCOPE OF THE ONTOLOGY

The following questions were considered to set the scope and limitations of the ontology:

- What is the domain that the ontology will cover?
  Representation of movies, actors and locations where movies are projected
- What are we going to use the ontology for?
  We plan to use the ontology for applications that suggest movie projection locations based on the combination of movie, movie type and actor

- For what type of question should the information in the ontology provide an answer?
  Where is the nearest Theatre that will project the movie "Shall We Dance?" at around 6PM?
  When and where can we watch a comedy movie?
- Who will use and maintain the ontology?
  Tourist-Agents and Agent-brokers

STEP 2: CONSIDER RE-USING EXISTING ONTOLOGIES
We will start our ontology from scratch. No existing ontologies will be used.

STEP 3: ENUMERATE IMPORTANT TERMS IN THE ONTOLOGY
Several concepts surrounding a movie event can be classified. Our classification will only cater for some of the major concepts illustrated in the following diagram in figure 3 below.

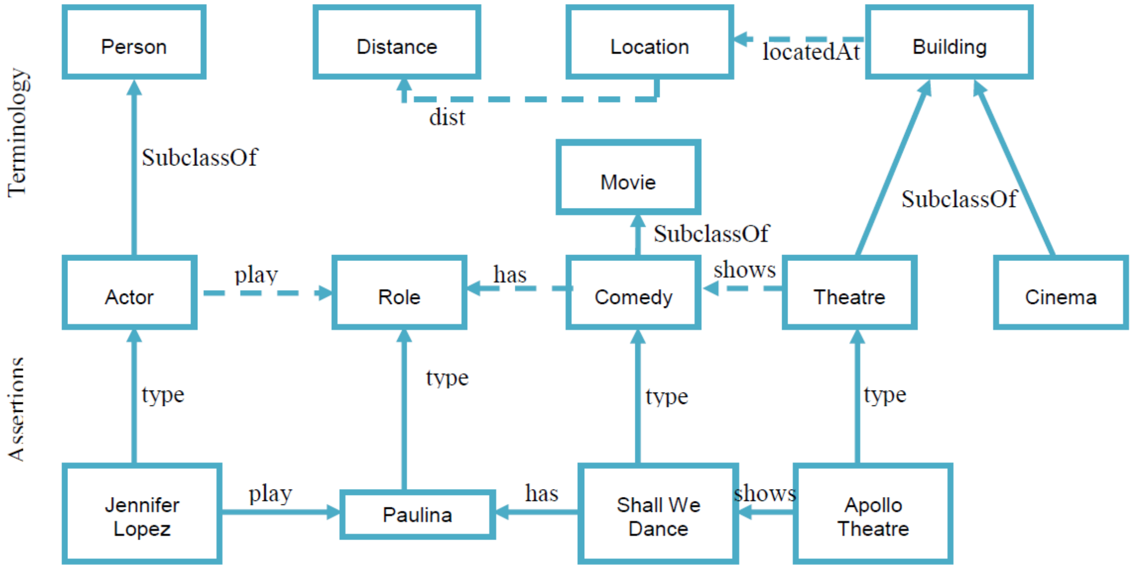

Figure 3. Diagram for Concepts classification

In figure 3, a comedy (e.g. "Shall We Dance?") – subclass of the "Movie" class is played in a theatre ("Apollo Theatre"). A theatre is a subclass of the "Building" class. The building class has a "locatedAt" property that provides the location of the building. This feature can be handy when looking for a theatre near a certain area. The Distance to or from the Theatre can be retrieved using the "dist" property.
A movie has several roles played by actors. For example, "Shall We Dance?" has the role of "Paulina" played by "Jennifer Lopez".

For reasons of space restrictions, we will not explain the rest of the process through which ontology is defined to facilitate communication and interaction between entities using OWL. Hence, the purpose of this exercise has been to demonstrate how the same procedure could be followed in building and establishing communication terminology for the multi-agent environment that the IDW model introduces. In the next section, we simulate an example of interaction between agents for a Tourism management System Scenario by building petrinets.

2.3. Theoretical Simulation of the IDW model: Tourism Case Study

2.3.1 Tourism Management System: Information Overload Context

An effective Tourism Management System is expected to, among other numerous operations, calculate the evolution of hotel rates, estimate the load and predict room availability within a specific period of time based on recurring patterns around the city, town, state or country for any visitor. The Data Warehouse should be able to serve in terms of decision making. The services expected from the application include a forecast of trends in the tourism industry at a given point in time, the competitors and fluctuations in offered services, the positioning of products in the minds of tourists (customers), customers' urgent needs, the rate of satisfaction etc. This strategic information ought to be provided as efficiently and effectively possible to keep the users and decision makers ahead of their competition. Furthermore, the role of travel intermediaries is not to be neglected. These include travel agents, tour operators, conference organizers and booking agents that make the use of computer reservation systems to service needy tourists. These applications are understood to be connected to data warehouses managing data around the world. These data span from transportation (air, water or surface) reservation details, accommodation (hotel chains, Motels, Gust houses…), and attraction outlets etc.

Moreover, these tourism companies around the world face any combination of input overload, where they must store and access data from operational systems, external databases, and third-party providers; output overload, in which data users want more and more information, tools, and analysis; and resource and administrative overload, which often happens when an organization outgrows its existing data warehouse and finds a replacement too costly to create.

Given tourism statistics, there are billions of records to be considered when one analyses the trends in the industry. All this analysis involves quite a massive volume of data that need to be processed and accessed to produced detailed analytical and statistical information needed for forecasting and service predictions. Numerous examples exist online that can be used to illustrate this fact. Let's consider a single example for Singapore.

"In 2010, Singapore attracted a record 11,638,663 visitors according to the Singapore Tourism Board's statistics, excluding Malaysian visitors who entered the country via the Causeway or the Second Link. This was a 20.2% increase over 2009.

Total visitor days was a record 44.9 million days, a growth of 17.1%, or an average of 3.86 days per visitor. 22.3% of visitors were day-trippers (less than one day), 19.7% stayed for one day, 17.6% stayed for two days, 14.7% for three days, and the rest for 4 days and above. 76.0% of visitors arrived by air, 10.9% by sea, and 13.2% by land. The largest age group of visitors was from 25–34 years old at 22.7% of visitors, followed by 21.8% for those aged 35–44 and 17.9% for those from 45–54 years old.

Over half (53%) of visitors came from the five biggest markets, namely Indonesia (2,305,149), People's Republic of China (1,171,337), Malaysia (1,036,918), Australia (880,486) and India (828,903). Other major markets include the Philippines (544,344), Japan (528,817), the United Kingdom (461,714), Thailand (430,022) and the United States of America (416,990).

Tourism receipts was estimated to reach S$18.8 billion in 2010, a growth of 49% compared to 2009, with Shopping and Sightseeing/Entertainment accounting for 21% of total expenditure each, Accommodation making up 19%, and Food and Beverage another 10%. Medical receipts, representing the medical-tourism industry the country was trying to promote, contributed 5%. In particular, Sightseeing/Entertainment, which included expenditure at the two new Integrated Resorts, grew by 1,834%.

Gazetted hotel room revenue was estimated at S$1.9 billion, an increase of 21.8% over 2009. The overall average occupancy rate was at 86%, 9.8% more than 2009, with the Economy tier seeing the largest increase of 15.6%. Overall average room rate was atS$212, an increase of 12.2%, while the overall revenue per available room was S$182, an increase of 26.6% over 2009.

General trends

Year	***Tourism Arrivals***	***Percentage change from previous period***
1965	*99,000*	
1970	*579,000*	*488.1%*
1975	*1,324,000*	*128.6%*
1980	*2,562,000*	*2.%*
1985	*3,031,000*	*18.3%*
1990	*5,323,000*	*75.6%*
1995	*7,137,000*	*34.1%*
2000	*7,691,399*	*7.76%*
2005	*8,943,029*	*16.27%*
2010	*11,638,663*	*30.14%*

." (Wikipedia contributors, 2013).

Intensive data processing occurred in acquiring this information. These data comprise of billions of records and flat files from over a period of time (even decades) to analyse trends and predict the performance of the Industry. The performance could be hindered due to information overload.

Before we simulate agents' collaboration, we look at few steps involved in building a Data Warehouse for a tourism management system. The following figure shows the complete Data Warehouse Bus Architecture. It helps us to think through how the Data Architecture will operate at run-time.It provides a common Framework for establishing consistent standards for data and structures. The top row shows a conceptual view that represents the major Building Blocks in the migration of data from the Operational Data Stores to the Semantic Layer that produces the detailed report needed for User View.

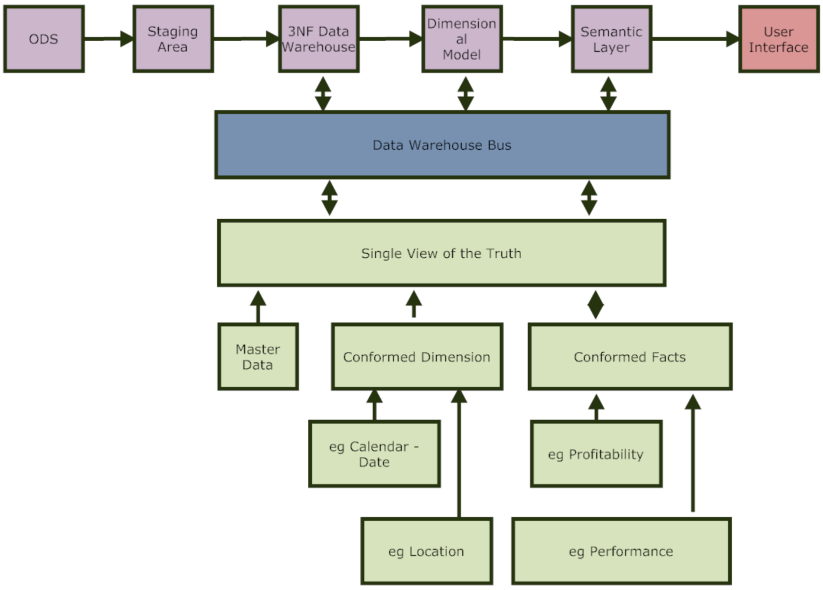

Figure 4. Data Warehouse Bus Architecture

In a data Warehouse, data are organised in dimension tables, facts tables, flat files etc. that are stored in data marts as shown in figure 8 later .

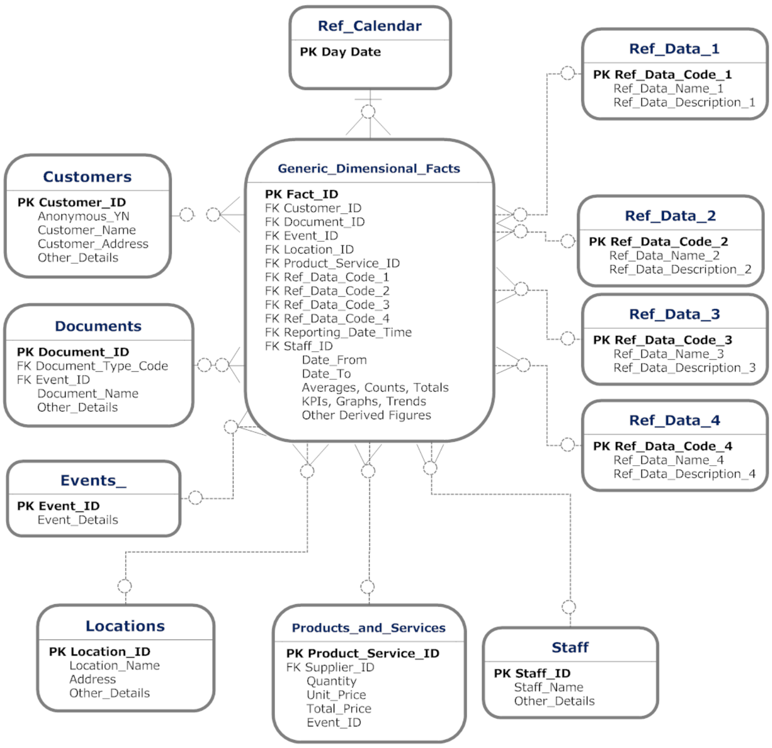

Figure 5. A generic Dimensional Model for Tourism Industry

This is a generic Dimensional Model which can be used to represent a range of events (such as check-in, order a tour package…), products, customers, locations, staff and all the tourism DW dimensions. Additionally, a typical tourism data Warehouse may now store documents, spreadsheets, images, RFID tags, sensor readings, videos, and a host of other formats. Governments, tourism agencies need to integrate any data so that analytics and Business Intelligence (BI) solutions can access information from throughout the organization to ensure complete, accurate analysis. Thus, a scenario is created where more cross-business users accessing more information, in more formats, from more data marts in a huge data warehouse.
Lower-priced storage has meant organizations can afford to collect all this information. But managing, analysing, and collaborating on this data is a different challenge. Below is a dimension table representation a check-in event for a hotel in a specific location at a given calendar date.

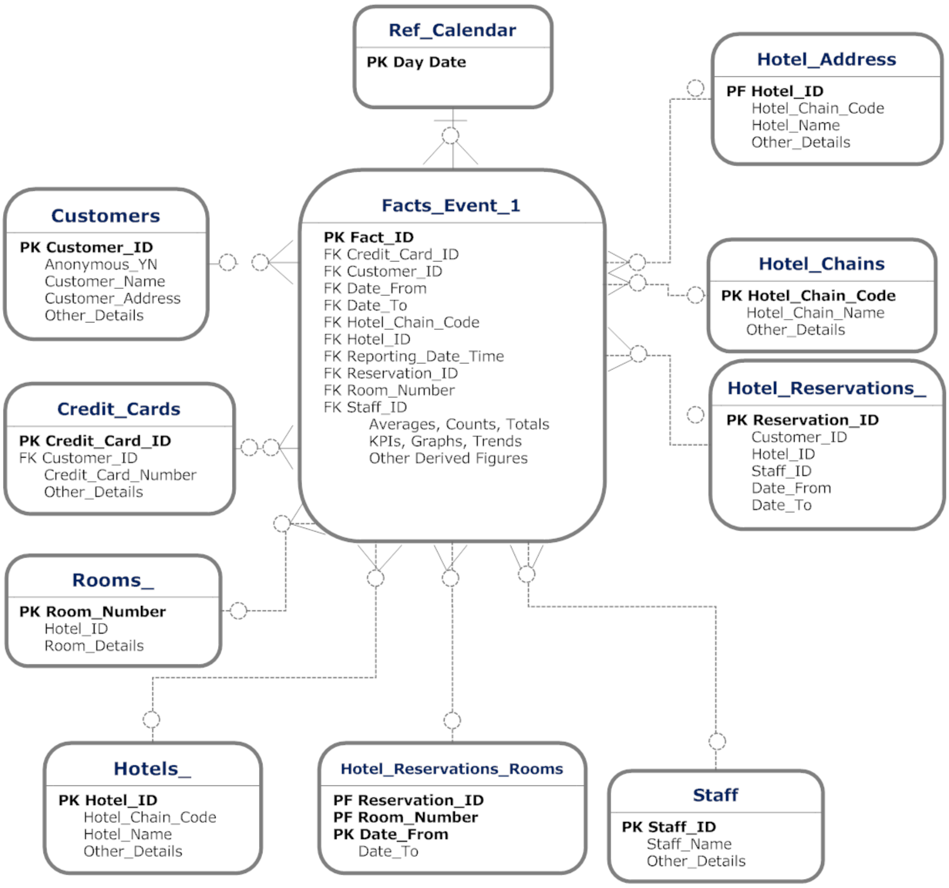

Figure 6. The dimension table representation a check-in event for a hotel

Dimensions to process the event include hotel staff, hotel chain, reservation room, rooms, payment methods, credit cards, addresses etc. with the corresponding flat table.
This can be explained graphically in the following Data model about the Hotel Check-In event (Fig.6). As in any data model, business rules are critical to understand and implement any transaction. The business rules involved in this case include:

• A Customer has one and only one Address.
• A Customer has one or more Addresses.
• A Hotel belongs to one and only one Hotel Chain.
• A Hotel has one and only one Address.
• A Reservation is associated with one Customer.
• A Reservation is associated with one member of Staff.
• A Room belongs to one and only one Hotel.
• A Room Card or Key is associated with one and only one Room and Reservation

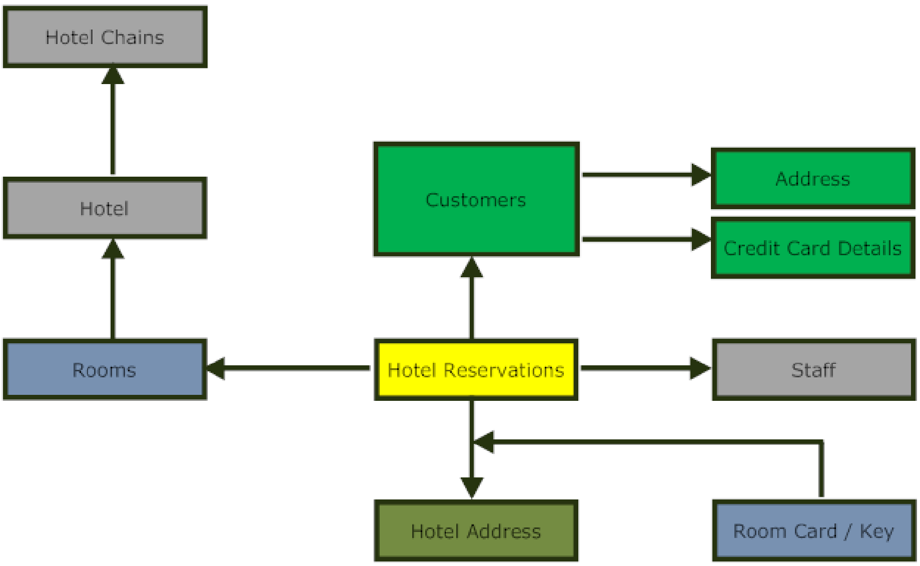

Figure 7. Data model about the Hotel Check-In event

In addition to these tables, separate Data Marts can be designed for Revenue, Costs and Complaints as shown in figure 7 below. A marketing Director for a given Hotel might need specific information in figures, such as how the Complaint levels are doing and how the profitability is affected by a promotions Campaign in a specific time-period.

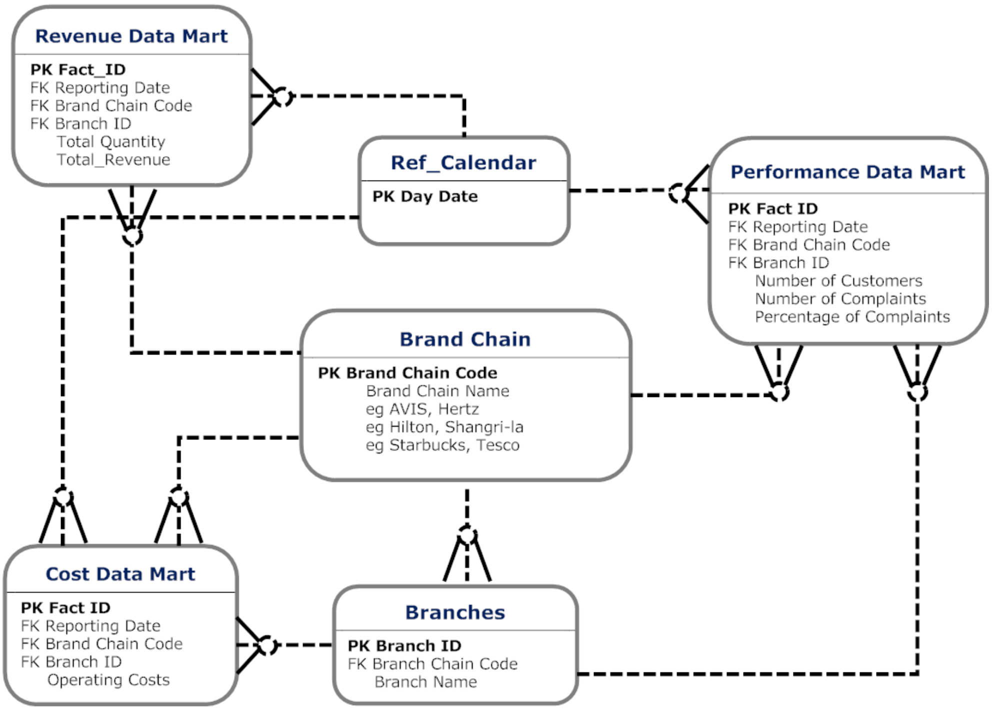

Figure 8. Data Marts can be designed for Revenue, Costs and Complaints

Additional data marts could be added for every event, product and process for thousands of clients and customers and this increases massive volumes of data. Given the volume of data to be processed and the complexity of technical tools to be manually used for this purpose, our model suggests the implementation of agent technology to alleviate through delegating these responsibilities. [6] emphasize that current OLAP tools provide good capabilities; however, these tools have proven difficult to use for common users. Even seasoned programmers and analysts find multidimensional analysis difficult. These tools undoubtedly add significant value to data analysis in data warehousing but are difficult to use for common user like a call center agent that needs to provide a list of alternative tour packages (Transportation, Accommodation, suitable Sports, Entertainment, etc.) for tourists in a short period of time according to the requester's budget.

Agents have the ability to acquire new knowledge based on the frequency of user actions and data access [3]. This is very critical as it saves to reduce the response time when Tourists enquire or request specific services that might have been provided prior as agents can easily reference and act speedily by using this knowledge. Their ability to learn allows for better performance as they can work independently. If an employee in a Tourism Agency makes use of an OLAP reporting tool, the result can be effectively and efficiently produced as mining Agents can facilitate the process. What the employee has to do is just to launch the tool and send a request and agents will process the rest.

The mobility of agents and their ability to move across multiple computers in a network can make data warehousing very efficient and effective. In Updating a data mart for example, agents can be very useful. As data marts are made up of several data sources, coordination will be significantly improved with agents as they will be able to coordinate these updates with load scheduling to see which source is updated and when, taking into account system use, historical load times, priorities and any other additional factors to be considered. In the next subsection, we simulate how collaboration between agents can help alleviate the information overload problem through tasks sharing and delegation.

2.3.2 Implementation of IDW model: Scenario

The scenario is about a tourism management system that interacts with a tourism data warehouse built on the IDW model framework. So many complex processes and events are involved in any such scenario to access billions of data records. However, in this scenario we consider a simple event with related processes to simulate the functionalities and operations of the IDW model. The processes involved in the scenario include: making a call (by the tourist) and agents attending to the call to help find and provide appropriate services to the caller. The core of the simulation is to illustrate the interactions between agents in finding required information in a simulative modelling manner.

The participants include:

a. Tourist
b. Call Centre agent
c. Service Provider Agent
d. Call Centre Service Management System

HIGH LEVEL INTERACTION DIAGRAM

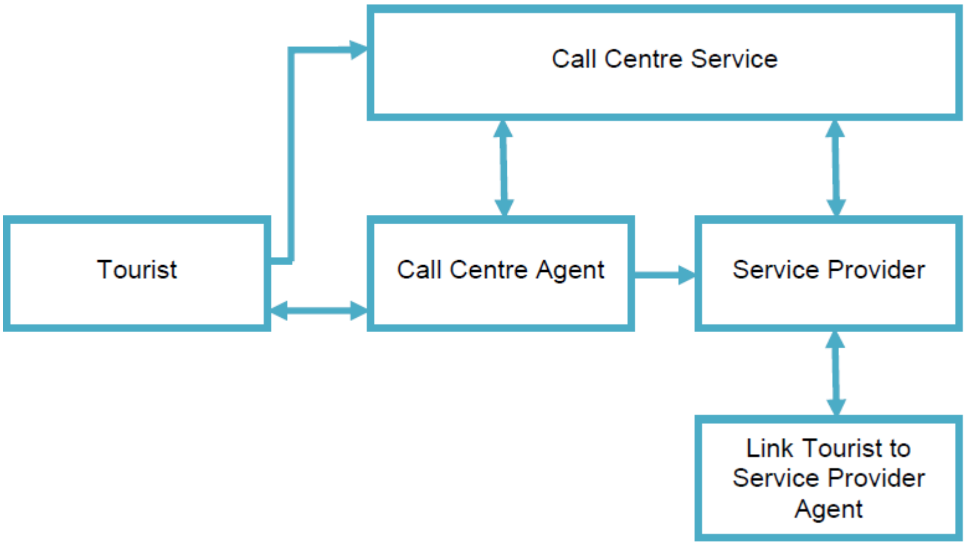

Figure 9. High-Level Interaction Diagram

Interaction between Agents

- A tourist will make a call to the Call Centre Service Management System (CCSMS).

- Upon receiving a call, the call centre Agent on the CCSMS will direct the call to the next available agent for processing.
- Then the communication will be established between the agent and a tourist in the form of service request.
- The Call Centre agent will check the requested service on the Service Management Module on the CCSMS.
- The call centre agent will link the tourist to the Service Provider agent with the regard to the service that needs to be provided.
- The Call Centre Service Management System (CCSMS) consists of Call Centre Agents and Service Management Agents.
- The Call Centre Agents will be responsible for managing the calls from the tourist to the call Centre Agent for directing and redirecting the calls.
- Then another set of Agents are Service Management Agents and they deal with managing, directing and checking the appropriate or suitable services to the Service Provider Agents.

CALL CENTRE CHAIN DIAGRAM

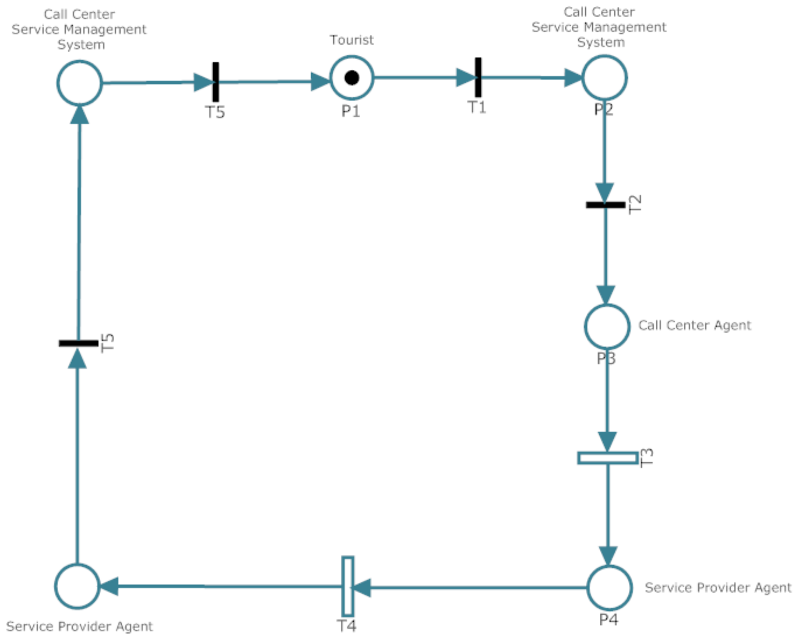

Figure 10. Call Centre Chain Diagram

T - Tasks performed by the participants
T1 - The tourist calls the Call Centre Service Management System (CCSMS)
T2 - The CCSMS links the tourist to the next available Call Centre Agent
T3 - This is an aggregation of services. The CCSMS finds the appropriate Service Provider Agent and links them with the tourist. T3 is detailed in the next figure.
T4 - The Service Provider agent will send feedback of the service to the Call Centre Service Management System
T5 - The Call Centre Service Management System terminates the transaction

CALL CENTRE AGENT DIAGRAM (T3 in the above call centre chain diagram)

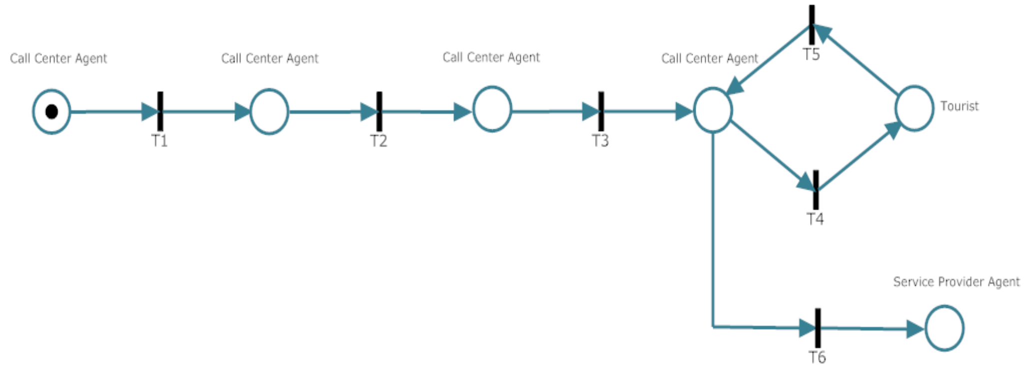

Figure 11. Call Centre Agent Diagram

T1 - Open a token into the system
T2 - Enquire about type of service the tourist wants
T3 - Search in the system for appropriate service provider agents
T4 - The Call Centre Agent proposes a list of service provider agents that meet the tourist's criteria
T5 - The tourist selects service provider of choice
T6 - The Call Centre Agent links the tourist with the selected service provider agent

SERVICE PROVIDER AGENT DIAGRAM (T4 in the above Call Centre Chain Diagram)

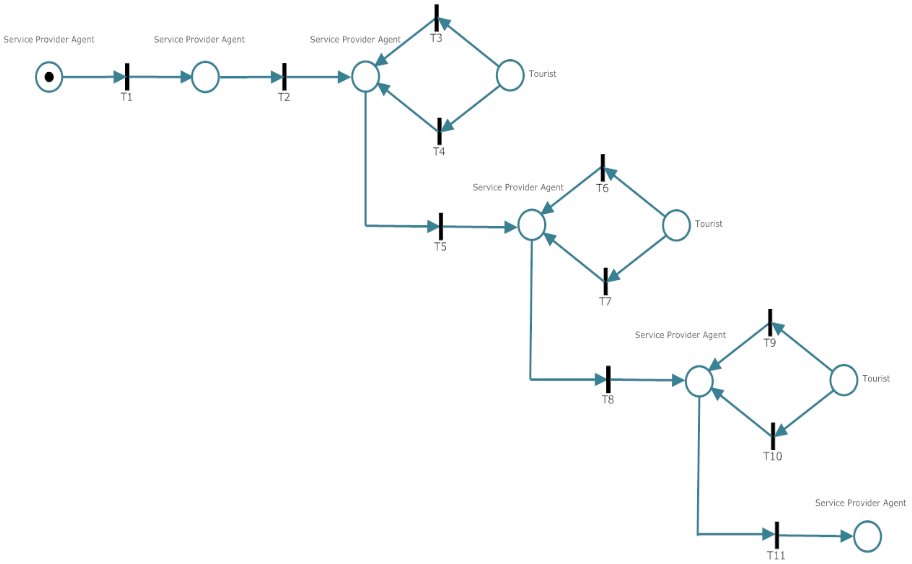

Figure 12. Service Provider Agent Diagram

T1 - Enquire further details of the requested service
T2 - List suggestions
T3 - Summary of the preferred accommodation
T4 - The tourist chooses the kind of accommodation he/she wants.
T5 - In case of transport-type of services, there will be several options to select.
T6 - List of all available transports
T7 - The tourist will select the sedan, two-seater type of transport
T8 - List all the kind of sports available
T9 - Summary of the preferred sports activities
T10 - Tourist can either select Golf, Swimming or others etc.
T11 - This occurs when the call is closed

The above scenario presents an ideal situation to be provided by any data warehouse designed using the IDW model presented in this mini-dissertation. As users (tourists) do not see the processing behind any application, it is of utmost importance to ensure that the requested information and services are provided with the optimal time. The scenario has presented an interactive collaboration between key actors in providing services to the of service requesters. The core of this interaction occurs between agents situated at different layers of the Data Storage and Processing component of the Data Warehouse. We support in this work that this scenario is possible through Agent technology as compared to manual data processing.

3. Conclusion

In this paper, we introduced and discussed the approach considered to tackle the effects of information overload on query response time in data warehousing. We discussed the intelligent data warehouse (IDW) model. We discussed important concepts to be considered for the implementation of the model. These concepts include the agent communication language, here referred to as Query Interchange language (QIL) and ontology, as well as the semantics. To demonstrate the applicability of these concepts, we considered two case studies to demonstrate the creation and use of ontology, as well as the interaction between agents using the tourism case study. The fundamental approach behind intelligent data warehouse is to integrate specialised features for the storage of data, the processing or cleaning of this data and, most importantly, the improved retrieval of this information.

Agent-integrated intelligent data warehouse is aimed at automating decision-related tasks in order to fully and successfully support the entire decision-making process. Thus, this operation will enable more effective data management, and reduce long time-consuming data access. Intelligent data warehouse provides improved procedures for the collection of relevant information, and generation of alternative decisions for given business problems. Intelligent data warehouses incorporate specialised functions to support flexible and adaptable data access from anywhere across the enterprise. In short, intelligent data warehouses offer the potential functionality to automate the entire problem- solving process that required complex mechanisms from traditional data repositories.Our case studies were a demonstration for simulative behaviour and operations of the IDW. We used petrinets to simulate the process of answer customers (tourists) needs upon receiving specific requests (queries).

While the research study has been conducted based on theoretical analysis and simulation of the model functionalities, the intent of this research project’s author is to practically implement the proposed approach with real data in order to test the discussed functionalities and approach for high-performance applicability. An empirical evaluation of the model will be at the centre of the author’s future research project. Currently, several programming languages for software agents design exist; some of these languages include C,C#,Tcl, Java, and TeleScript etc. We plan to examine and assess any available programming languages for the implementation of this research. Tools and libraries such as the Object Request Broker Voyager can be considered as it allows agents movements and communication in Java. In order to access data sources, protocols such as the Open Database Connectivity (ODBC) driver or the Java Database Connectivity (JDBC) among others can be considered.

Thus, the ambition of this paper has not been to create the actual tool (IDW) but rather to develop the model and show its applicability. As mentioned above, this work is a stepping stone for computer scientists (especially data engineers) to focus on the next generation of Data Warehouses, which we referred to as Intelligent Data Warehouses.

References

[1] Albayrak, S., Wollny, S., Varone, N., Lommatzsch, A., & Milosevic, D. (2005, March). Agent technology for personalized information filtering: the PIA-system. In Proceedings of the 2005 ACM symposium on Applied computing (pp. 54-59). ACM.
[2] Castellini, P., Cristalli, C., Foehr, M., Leitão, P., Paone, N., Schjolberg, I., & Wagner, T. (2011, November). Towards the integration of process and quality control using multi-agent technology. In IECON 2011-37th Annual Conference on IEEE Industrial Electronics Society (pp. 421-426). IEEE
[3] FASLI M., 2007. *Agent Technology for e-Commerce*. England: John Wiley & Sons, Ltd
[4] Färber, F., Cha, S. K., Primsch, J., Bornhövd, C., Sigg, S., & Lehner, W. (2012). SAP HANA database: data management for modern business applications. ACM Sigmod Record, 40(4), 45-51
[5] Ginmen, M. (2012). Information culture and business performance.
[6] Glick B., 2008. *How to improve business performance with data warehouse technology*. London : VNU Business Publications Ltd
[7] Hemp, P. (2009). Death by information overload. Harvard business review, 87(9), 83-89.
[8] Labrou, Y., Finin, T., & Peng, Y. (1999). *Agent communication languages: The current landscape*. Intelligent Systems and Their Applications, IEEE, 14(2), 45-52.

[9] Onal, E., & Endsley, M. R. (2012, September). Requirements and Design for Better Cultural Situation Awareness: Delivering the Right Information. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 56, No. 1, pp. 228-232). SAGE Publications
[10] Rahimian, F., Girdzijauskas, S., Payberah, A. H., & Haridi, S. (2012). Living In the Information Era. In the 32nd International Conference on Distributed Computing Systems (ICDCS 2012).
[11] Rob P., Coronel C., Crockett K., 2008. *Database Systems: Design, Implementation & Management.* United Kingdom: Cengage Learning EMEA. International Edition
[12] Vijayan S, Ranjit B.,1999. *Data analysis and mining environment: a distributed intelligent agent technology application*. USA
[13] Wikipedia contributors (2013). Tourism in Singapore. Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia, 5 Feb. 2013.
[14] Zhang, D. L., & LI, H. Z. (2011). On security issues in cloud computing. Hunan Agricultural Machinery, 5, 046.

BIOGRAPHIES OF AUTHORS

Patrick Mukala
He holds a master in software engineering from the Tshwane University of Technology, South Africa. His research interests span from Software and Data Engineering to complex networks theory, information processing and Semantic Web. He is on the verge of completing a second master degree from the same institution.

Prof. Van der Walt
He is a Professor at the Tshwane University of Technology, now retired preofessor, He still provides his experience and assistance in supervising and guiding postgraduate students from multiple spectrum of computer science.