Business Intelligence Architecture Informed by Organisational Semiotics

John Effah, Prince Senyo, Stephen Opoku-Anokye

Business Intelligence Architecture Informed by Organisational Semiotics

John Effah[1], Prince Kwame Senyo[1,2] [0000-0001-7126-3826] and Stephen Opoku-Anokye[2]

[1] Department of Operations and Management Information Systems, University of Ghana Business School, Ghana,

[2] Informatics Research Centre, Henley Business School, University of Reading, Reading, United Kingdom

jeffah,pksenyo{@ug.edu.gh},stephen.opoku-anokye@henley.ac.uk

Abstract. This study draws on organisational semiotics and design science methodology informed by abductive reasoning to develop a business intelligence (BI) architecture. Organisational semiotics research has so far paid limited attention to BI in general and its architecture in particular. Moreover, BI research in information systems (IS) focuses largely on either technical or social activities. Organisational semiotics offers frameworks and model which can be used to develop a BI architecture with combined technical and social views. This study therefore develops a BI architecture based on knowledge hierarchy, semiotic framework, and semiotic activity hierarchy. The paper uses a manufacturing company's BI experience as a case study to inform and evaluate the proposed architecture. The study's contribution stems from its development of the organisational semiotics informed BI architecture and its implications for research and practice.

Keywords: Organisational Semiotics, Business Intelligence, Knowledge Hierarchy, Forma-Informa-Performa, Semiotic Framework First Section

1 Introduction

Following technological advancement in data infrastructure as well as tools and techniques for analytics and data mining, business intelligence (BI) has increasingly attracted research attention in information systems [5]. Recent conceptualizations view BI as the ability to acquire and apply actionable knowledge to make decisions [8]. As a multidisciplinary concept, BI lacks a consensual definition [7, 16]. However, in information systems, BI refers to the use of information and communication technologies (ICTs) for data gathering and storage to generate actionable knowledge for decision making [15, 17]. BI has generally been viewed as a process [15]. However, this study argues that beyond being a process, BI can also be an output as actionable knowledge.

Thus far, organisational semiotics research on BI remains limited. The few studies on the subject have focused largely on data visualization [11, 13], which is only a component of the final stage of the BI process. As a result, not much is known about BI process in organisational semiotics. A recent study on information architecture [21] makes a case for increasing the scope of BI research in organisational semiotics. The current study responds to such a call by developing a semiotic informed BI architecture.

The rest of the paper is structured as follows. Section 2 reviews related works on knowledge hierarchy and BI systems. Section 3 presents semiotic framework and semiotic activity hierarchy as the study's theoretical foundation. Section 4 presents BI experience of a manufacturing firm's human resource intelligence system as a case for developing and evaluating the BI architecture. Section 5 develops the BI architecture based on the related works, the theoretical foundation and the case study through an iterative process. Section 6 discusses the results of the study. Finally, Section 7 provides the conclusion with suggestions for future research.

2 Related Works

2.1 The Knowledge Hierarchy

Knowledge hierarchy [1, 18], also called information hierarchy or wisdom hierarchy, represents the structural and functional relationships between data, information, knowledge and sometimes wisdom [4, 18]. Figure 1 shows the knowledge hierarchy and its layers.

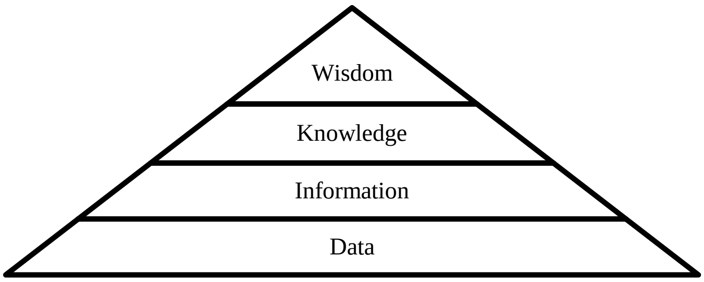

Fig. 1. Knowledge Hierarchy [18]

Like knowledge itself, concepts within the hierarchy lack consensual definitions across disciplines. Table 1 however offers generic definitions as used in information systems [e.g., 4, 18]. Data refers to symbolic facts captured and stored in media; information constitutes statements with meanings; knowledge refers to true statements that are socially believed and verified; while wisdom refers to applied knowledge judged to be right and socially acceptable.

The functional perspective shows the hierarchy as dynamic interactions between the elements. Thus, data undergoes processing to derive information, which is analyzed to generate knowledge, which is judged or assessed to get wisdom. Knowledge hierarchy has been applied with semiotic framework in organisational semiotics research [e.g., 2]. However, its relationship with BI remains limited. The current study therefore adapts the knowledge hierarchy to derive an intelligence hierarchy as part of the process for developing the BI architecture.

Table 1. Definitions of knowledge hierarchy concepts

Element	Meaning
Wisdom	judgements that are socially desirable
Knowledge	beliefs that have been socially verified to be true
Information	meanings derived from processed data
Data	Symbolic facts that have been captured

2.2 Business Intelligence Systems

Business intelligence systems [17] refer to a collection of technologies and techniques for capturing, preparing and transforming data into knowledge for decisions. Figure 2 shows a framework for BI system and related components.

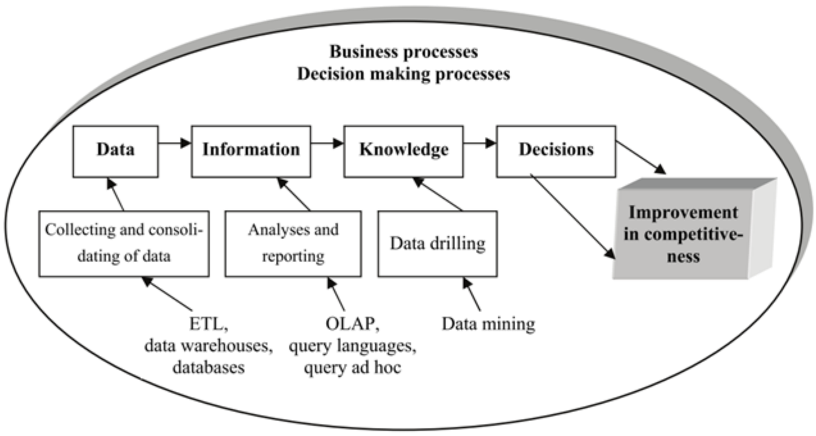

Fig. 2. Business intelligence system [17]

Figure 2 presents four stages of the BI system and an embedded three-layer model. The stages comprise data, information, knowledge and decisions. Each stage has underlying techniques supported by relevant tools. Thus, first data collection and consolidation techniques use ETL, data warehouse and database tools to generate data. Second, data analyses and reporting techniques depend on OLAP and query tools to generate information. Finally, data drilling uses data mining tools to generate knowledge to support decisions for improved business processes and competitiveness.

A key limitation of the intelligence system is the failure to conceptualise and incorporate intelligence as a fundamental component of the model. Thus, the model uses BI only as a process and not as an output that emerges at some point of the intelligence process. In simple terms, not all knowledge may be useful for decision making. To address this limitation, this study draws on the BI system to develop the organisational semiotics informed BI architecture in Section 4.2.

3 Organisational Semiotics

Organisational semiotics draws on signs to study information and communication systems in organisational context [10, 12]. A sign refers to whatever that stands to someone for something [9]. The current study draws on the semiotic frameworks and semiotic activity hierarchy as a combined theoretical foundation for the BI architecture.

3.1 Semiotic Framework

The semiotic framework (also called semiotic ladder) hierarchically structures sign systems into technical and social layers as shown in Figure 3.

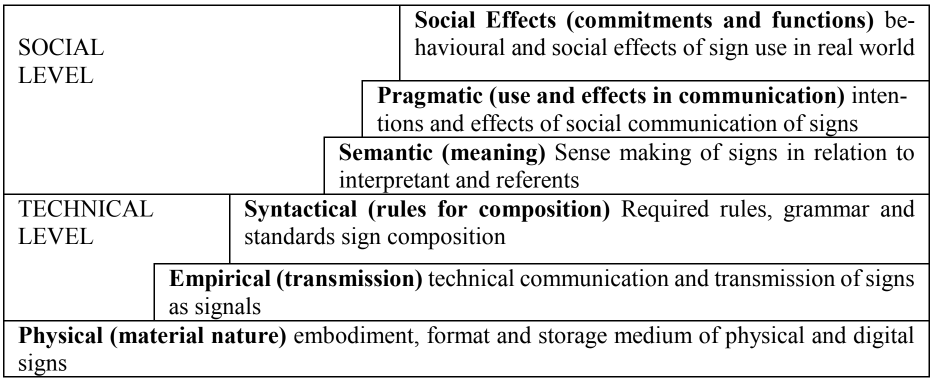

Fig. 3. Generic semiotic framework: Adapted from [14]

The technical layers comprise the physical, empirical and syntactic. First, the physical constitutes the material and digital components of signs. Second, the empirical concerns observable properties of signals as signs in transmission through a communication medium such as speed, capacity, efficiency and errors. Third, the syntactic concerns rules and standards regarding the physical composition and structure of a sign.

The social layers comprise the semantic, pragmatic and social effects. The semantic deals with meanings that signs convey. The pragmatic relates to the intentions as well as use and effects of signs in communication [10]. The social concerns change that communication and use of signs effect in the real world. Such effects include changes in the status quo that results from social activities such as agreements, norms and decisions. The next section presents semiotic activity hierarchy based on the semiotic framework.

3.2 Semiotic Activity Hierarchy

The notion of semiotic activity is introduced in this study as layers of formative, informative and performative activities based on the semiotic framework as shown in Figure 3. The formative activity involves the physical, empirical and syntactic layers for composition and transmission of signs as data; informative activity relates to interpretation and communication of signs as information; finally, performative activity informs the use of actionable knowledge to make a decision that effect changes in the real world.

SEMIOTIC ACTIVITES	**Performative Stage –** activities based on actionable knowledge for decision making and generating social effects
	Informative Stage – activities and outputs based on semantic and pragmatic layers involving information use in communication
	Formative Stage: activities and outputs physical, empirics and syntactic layers involving data

Fig. 4. Semiotic Activities Hierarchy

The layers of activities are based on the notions of forma, informa and performa as patterns of signs systems developed by Dietz [6] and related to the semiotic framework by Beynon-Davies [3] as semiotic acts. According to Dietz [6] forma refers to material or physical composition of a sign; informa deals with content and meaning of signs; while performa deals with communication and use of signs for making decisions for social actions.

4 Case Study

This section presents the BI experience of a multi-site manufacturing company in the UK, MSMC (pseudonym) involving human resource and diseases data warehousing and mining. In 2010, the company implemented a BI solution with a data warehouse for consolidating and restructuring operational data including that of human resources; metadata layer for providing meaningful data views from the data warehouse and marts; and presentation layer for reporting and analytics, including pre-built reports, ad-hoc queries and analysis as well as BI visualization.

In 2011, the Human Resource (HR) Director had a requirement to know monthly trends of sickness among the workforce across all sites. To do this, the BI team extended the existing data warehouse to include disease outbreak data in areas of the various sites. Subsequent data mining activities established associations between the health data and music festival data. Hence, additional data on music festivals were incorporated into the data warehouse and the HR data mart. The following section shows the semiotic activities that occurred at various stages of the BI architecture.

4.1 Formative Activities

The physical implementation of the data store layer of the BI solution included a consolidation and restructuring data warehouse (CRDW). The techniques used at the formative stage involved the use of extract, transform and load (ETL) tools to collect data from operational and external sources including HR data, transforming the data through restructuring and cleaning and loading data into the warehouse.

The HRM system primarily supported the operations of the HR Department. As such, it was developed to capture and store employee related data across the organisation, including data on sickness and sick leave. The organisation's policy on sick leave allows employees to self-certify if the sickness period is not more than 7 days in a single period or 10 days in total for a whole year. Within these periods, self-certified sick leave does not require a note from a medical practitioner or evidence of visit to a medical centre. As part of the formative activities, data were collected by the recording of sick leave taken using the organisation's Human Resource Management (HRM) System. Also, the data were extracted from the HRM System and loaded into an Organisation-wide integrated data warehouse. This was done to consolidate data on sick leave and from other sources, and to prepare the data for reporting and analytics.

Beyond the internal sources, ETL tools were used to capture and load NHS data on disease outbreaks and music festivals related to areas close various sites of the company. In sum, the BI formative activities made data available for the informative and performative activities.

4.2 Informative Activity

At the informative stage, a metadata set was setup with a prebuilt report for the monthly sickness trend to meet the information requirements of the HR director. In January 2012, the monthly trends report up to December 2011 highlighted August as the month with over 50% of recorded employee sickness, which had been consistent for the past 6 years. A further analysis of SiteX (pseudonym) of the company highlighted the 3rd week of August as accounting for over 75% of all reported employee sicknesses across the same 6-year period. Moreover, over 85% of all reported sickness in August for the 6-year period were self-certified, of which 90% were from one particular manufacturing site.

4.3 Performative Activity

Beyond the employee sickness data, health data from the National Health Service were analysed to identify any possible disease outbreaks in August, especially, the 3rd week. However, none was found. Analysis of data from the two water companies for the residential areas within commutable distance or 2-hour drive from SiteX also revealed no reported contamination in the periods concerned.

However, a search of local activities in the immediate surroundings of SiteX identified a music festival that starts on the 2nd weekend of August. Further investigations showed the festival has been running for over 50 years and has become very popular over the last 15 years. Further analysis revealed that over 70% of the people who attend this music festival as between the ages of 20 and 37 years.

Combining this external data from the organisers of the music festival with the internal HR data. Mining of the combined data revealed a strong correlation between the festival start date and the sick leave taken over the last 15 years. Also, the data revealed an exponential increase of sick leave taken during the 3rd week in August from 15 years ago to 6 years ago where it plateaued. Armed with this information, the senior management team decided to change the policy on self-certified sick leave from 7 days to 2 days in any single period.

5 Business Intelligence Architecture

This section presents the organisational semiotic informed BI architecture, which was developed through design science methodology informed by abductive reasoning [20]. The process began with the third author's observation of challenges that were associated with lack of a BI model to support the case organisation's decision and policy evaluation on sick leave. Following this, the study drew on the conventional knowledge hierarchy to develop an intelligence hierarchy as shown in Figure 5.

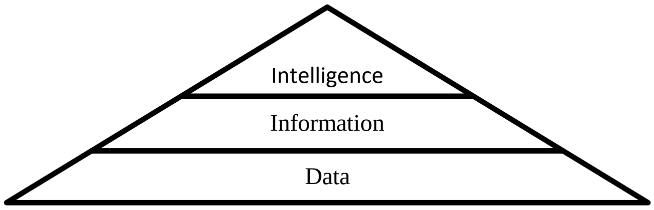

Fig. 5. Intelligence Hierarchy

As Figure 5 shows, the notion of intelligence replaces knowledge and wisdom. In this study, intelligence is considered as a form of knowledge that supports decision making and therefore actionable; hence the decision to use it to substitute knowledge and wisdom. In relation to the semiotic activity hierarchy, formative activity generates data, informative activity generates information while performative activity in the form of data mining generates intelligence to support decision making.

Subsequently, the study drew on the intelligence hierarchy and the semiotic framework with activity hierarchy to develop an initial BI architecture, which was validated and refined with the case study. Thus, the BI architecture process followed an iterative process whereby the BI experience in the case study served as a guide and evaluation criteria for refinement. Figure 6 shows the final architecture that emerged from the iterative process.

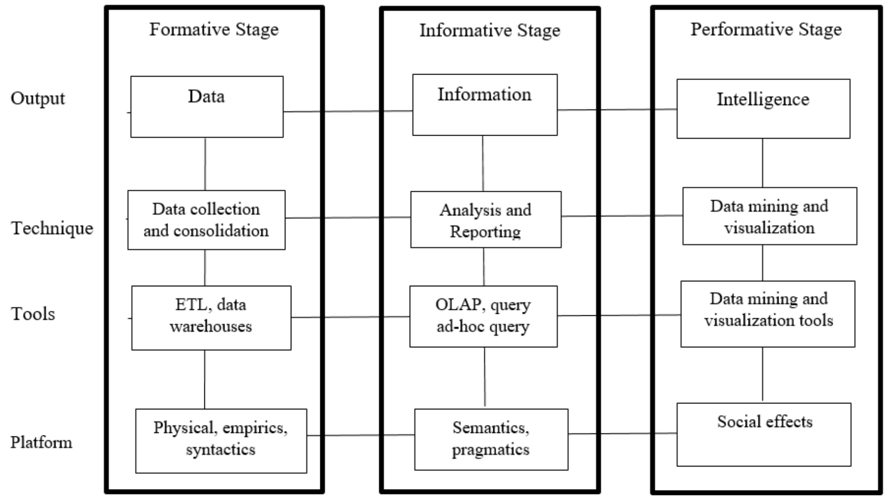

Fig. 6. Business Intelligence architecture

The architecture shows 4 layers and 3 stages with their output and supporting activities, techniques and tools. The interconnecting lines show how the various components are intertwined to generate the required intelligence to support decision making for organisational activities. The intelligence generation process begins from the formative stage where the platform layer relies on the tool and technique layers to generate data as output. This stage supports the informative stage to generate information as output. Finally, the performative stage generates intelligence as output, in the form of actionable knowledge for organisational decision making

6 Discussion

The case study shows the three stages of BI process and output as well as supporting techniques and tools of the proposed architecture. The formative activity corresponds to the data warehouse and ETL as tools, extraction and consolidation of internal as well as external data as techniques and the accumulated data as the output. In relation to the semiotic framework, the formative activity relies on the technical platform: physical, empirics and syntax.

The informative activity relates to metadata and reporting of monthly trend reports on sick leave. The underlying technique was analysis and reporting with OLAP, general query and ad-hoc query tools as supporting technologies. In terms of the semiotic framework, the informative activity relates to semantics and pragmatics in terms of sense-making and communication of information.

Finally, the performative activity involved the use of intelligence knowledge to make a decision by changing the policy on sick leave. The supporting technique involve data mining and data visualisation while the underlying tools were data mining and visualization software. In relation to the semiotic ladder, the performative activity relies on the social world effects.

The three activities show a clear distinction between data, information and intelligence in the architecture. Data is considered as symbolic facts that are accumulated to support informative and performative activities. Informative activity produces meaningful messages to provide answers to known questions through analytics as in the case of the HR director. Intelligence however presents actionable knowledge based on unknown and unexpected patterns, relationships, and associations as was the case with the sick leave and the music festival period. However, in practice, the three are not independent but highly related as demonstrated by the various intersecting line of the architecture.

Existing BI studies [e.g., 15, 19] largely portray the concept from a technical perspective and pays less attention to the social dimension. Also related organisational semiotic studies discuss some part of the subject in relation to knowledge management [e.g., 2] but does not link to BI. Our study therefore comes as the first to use organisational semiotics to develop a BI architecture with clear distinction between data, information and intelligence and their relationships at the semiotic, tools, techniques output and activity levels.

7 Conclusion

The purpose of the study was to develop a BI architecture based on the organisational semiotics framework. The study therefore presents a BI architecture founded on the organisational semiotic framework, intelligence hierarchy, semiotic activity framework based on formative, informative, and performative activity levels. Viewing intelligence as an actionable knowledge for decision making, the study contributes to organisational semiotics research by extending it to the domain of BI. It also contributes to BI research by basing the architecture on semiotic principles and frameworks.

For contribution to practice, the findings present a clearer BI process that intelligence analysts, developers and users can draw on to identify relevant technologies, techniques and activities that are required to develop and deploy a BI system in an organisational setting. In addition, the architecture presents a clearer network between data, information and intelligence to inform practices on how to develop such a system.

The limitation of the study stems from its exploratory nature and single case illustration in a human resource intelligence system. Given that BI does not focus on a single domain or problem area as does decision support systems, future research will evaluate the architecture in a multi-domain environment such as supply chain and customer relationship management

References

1. Ackoff, R.L.: From data to wisdom. J. Appl. Syst. Anal. 16, 1, 3–9 (1989).

2. Baskarada, S., Koronios, A.: Data, Information, Knowledge, Wisdom (DIKW): A Semiotic Theoretical and Empirical Exploration of the Hierarchy and its Quality Dimension. Australas. J. Inf. Syst. 18, 1, 5–24 (2013).
3. Beynon-Davies, P.: Formatics. In: Americas Conference on Information Systems. , Seattle, Washington, August 9-12, 2012 (2012).
4. Boell, S.K., Cecez-Kecmanovic, D.: What is "Information" Beyond a Definition? In: Thirty Sixth International Conference on Information Systems. pp. 1–20 , Forth Worth (2015).
5. Chen, H., Storey, V.C.: Business Intelligence and Analytics: From Big Data to Big Impact. Mis Q. 36, 4, 1165–1188 (2012).
6. Dietz, J.: The deep structure of business processes. Commun. ACM. 49, 5, 59–64 (2006).
7. Grossmann, W., Rinderle-Ma, S.: Fundamentals of Business Intelligence. Springer, Berlin Heidelberg (2015).
8. Leismeister, J.: Collective Intelligence. Bus. Inf. Syst. Eng. 4, 2, 245–248 (2010).
9. Liu, K.: Requirements reengineering from legacy information systems using semiotic techniques. Syst Signs Actions Int J Commun Inf Technol Work. 1, 1, 38–61 (2005).
10. Liu, K.: Semiotics In Information Systems Engineering. Cambridge University Press, Cambridge (2000).
11. Liu, K.: Semiotics in visualisation. In: 2014 9th International Conference on Evaluation of Novel Approaches to Software Engineering (ENASE). pp. 1–3 (2014).
12. Liu, K., Li, W.: Organisational Semiotics for Business Informatics. Routeledge, London (2015).
13. Liu, K., Tan, C.: Semiotics in Digital Visualisation. 3–13 (2015).
14. Mingers, J.: Guidelines for Conducting Semiotic Research in Information Systems, (2014).
15. Negash, S.: Business intelligence. Commun. Assoc. Inf. Syst. 13, 177–195 (2004).
16. Olszak, C.: Toward Better Understanding and Use of Business Intelligence in Organizations. Inf. Syst. Manag. 33, 2, 105–123 (2016).
17. Olszak, C., Ziemba, E.: Approach to building and implementing Business Intelligence systems. Interdiscip. J. Information, Knowledge, Manag. 2, 135–148 (2007).
18. Rowley, J.: The wisdom hierarchy: representations of the DIKW hierarchy. J. Inf. Sci. 33, 2, 163–180 (2007).
19. Sahay, B.S., Ranjan, J.: Real time business intelligence in supply chain analytics. Inf. Manag. Comput. Secur. 16, 1, 28–48 (2008).
20. Tan, C. et al.: Information Architecture for Healthcare Organizations: the case of a NHS Hospital in UK. In: ICIS 2013 Proceedings. pp. 1–11 (2013).
21. Tan, C., Liu, K.: An organisational semiotics inspired information architecture: Pervasive healthcare as a case study. ICISO 2013 - Proc. 14th Int. Conf. Informatics Semiot. Organ. IFIP WG8.1 Work. Conf. 35–44 (2013).