Open Data for Competitive Intelligence at a Mining Company: a Focus Group Approach

Vitor Afonso Pinto[1], Fabricio Ziviani[1], and Fernando Silva Parreiras[1]

LAIS – Laboratory for Advanced Information Systems, FUMEC University
Av. Afonso Pena, 3880, Belo Horizonte 30130-009, Brazil
`vitor.afonso.pinto@gmail.com`, `fabricio.ziviani@fumec.com`,
`fernando.parreiras@fumec.br`

Abstract. Competitive Intelligence is a process which involves retrieving, analyzing and packaging information to offer a final product that responds to the intelligence needs of a particular decision maker or community of decision makers. Competitive Intelligence Professionals transform raw data and information into intelligence by collecting and organizing information resources. Mining companies have a Mineral Exploration Department in charge of a comprehensive geological research program all over the world focused on the discovery of mineral deposits. This paper presents a proposal for extracting the information and the data sources commonly used by competitive intelligence professionals responsible for analyzing the mining market. Specifically, this paper intends to answer the following research question: What is the information used to analyze external environments, from the mining companies perspective? To achieve this goal and taking the social-technical design into account, we performed a focus group with experienced professionals from a mining company. As a result of this work, we present: 1) a list of generic data sources; 2) a categorized list of common information used to analyze the external environment; and 3) the main difficulties reported by professionals of this area.

1 INTRODUCTION

Organizations perform coordinated actions to seek, treat, distribute and protect information in order to obtain the sufficient knowledge to understand both the internal and external environment as a whole [25]. The set of activities performed by organizations to gather information about markets, competitors and products is called Competitive Intelligence [17].

Linked Open Data initiative makes lots of datasets available for anyone to use. According to [23], 90% of information required for understanding competitors and markets are public and available for everyone. Besides that, organizations are integrating data generated both inside and outside their boundaries, intending to leverage their competitive intelligence process [22].

Mining companies have their organic growth related to their competitive intelligence process. While geologists conduct researches intending to discover

potential areas for mineral exploration, competitive intelligence specialists analyze the feasibility of extraction, beneficiation and transportation of ores, taking into account technical, economic, environmental and social features.

Although everyone knows the relevance of asking the right questions in the context of competitive intelligence, people in charge of those activities have different academic background, such as: business administration, economics, accounting, engineering, law, among others. All disciplines must work in an integrated way so as to create long-term value and provide benefits to the company. Nevertheless, this may create a challenge as the minimum set of information may be different just because every professional in charge of each particular analysis may have a single perspective.

A focus group is a method that highlights the advantage of working in teams, which is a principle covered by social-technical design. This study presents how a focus group was used to raise requirements for a competitive intelligence software which should be based on open data and focused on mining companies. Specifically, this paper adresses the following research question: From the mining industry perspective, what is the information used to analyze external environments? A set of information used by competitive intelligence professionals from mining industry is presented, with the purpose of helping those people involved in competitive intelligence process to understand the minimum required to create intelligence packages.

Similar approaches for raising information used to generate intelligence packages for the real world are rare because companies prevent the sharing of any information with potential to create any kind of competitive disadvantages. However, the impressions presented in this paper were extracted from data collected through a focus group whose main purpose was to identify the information processed in order to assess potential places for investing.

This paper is structured as follows: Section 2 describes the main concepts used through the paper. Section 3 presents the approach used for realizing the focus group and Sections 4 and 5 describes the results and their discussion. We analyze the related work in Section 6 and conclude the paper with Section 7.

2 BACKGROUND

2.1 Competitive Intelligence

Intelligence gathering goes on every day, without necessarily being called by its rightful name. It implies legal research efforts made by businesses studying their competitor's products, organizations and related matters. As this gathered intelligence is used to help business positioning, it is called **Competitive Intelligence (CI)**. CI is defined as the use of public sources to develop information on competition, competitors, and the market environment, including economic, regulatory, political, demographic influences, etc. [6].

Competitive Intelligence runs the gamut of corporate goals, strengths, weaknesses, personnel, product and market entry plans and strategies. In other words,

48

Competitive Intelligence embraces how competitors have behaved, how they are behaving, and in a given environment how they are likely to behave [16]. It also includes information collected on actors and situations relevant to a competitive landscape, such as customers, suppliers, and relevant technologies [14].

Competitive Intelligence aims to monitor an organizations external environment for information relevant to its decision-making process. Companies that have formal and well-organized competitive intelligence units enable managers to make informed decisions about critical business matters such as investment, marketing, and strategic planning [4].

Competitive Intelligence is, at the same time, both a Process and a Product. It is a Process because it is a systematic and ethical program for gathering, analyzing, and managing external information that can affect the companys plans, decisions, and operations. It is also a Product because it is essential to have an application that can handle formats such as text, tables, graphs, pie charts, photographs, drawings, or multimedia files with sound, in order to convey intelligence in the ways CI requires [3].

To support CI as a process, there must be corporate mechanisms for following up on such information that trigger the initiation of a structured search process to complete and update information from individuals [7]. In this perspective, CI Professionals transform raw data and information into intelligence by collecting and organizing information resources, a fundamental step, but also by extracting information and adding value by assessing the relevance of the information in a particular context [3].

2.2   Semantic Web

In order to perform Competitive Intelligence, companies need to integrate internal and external data. In this context, Semantic Web is presented as a solution because it facilitates interlinking of heterogeneous data sources. The idea of Semantic Web is to have data on the web defined and linked in a way that it can be used by machines - not just for display purposes, but for using it in various applications. [28]. It also carries the promise to make web machine-understandable by enriching available information with logic-based semantics and provide us with a new paradigm for knowledge interchange and sharing. [5].

Linked Open Data   A pragmatic vision of the Semantic Web has emerged via the Linking Open Data project (LOD), focusing on translating data sets available on the Web into RDF, which is a directed, labeled graph data format for representing information on the Semantic Web. Data sets have been provided via this LOD initiative, such as DBpedia (the RDF export of Wikipedia) or Geonames (a large geolocation database). They form a complete Web scaled graph of interlinked knowledge, known as the Linked Open Data Cloud [21]. Open data is data available for anyone to use and re-publish. An open data environment is essential if the Semantic Web is to provide a common framework that allows data to be shared and reused across application, enterprise, and community boundaries. [27]

Linked Enterprise Data According to [8], another goal of Semantic Web is to use the Semantic Web technologies for Enterprise Information Integration, by considering all the data of an enterprise, or of a closed community, as one global database, creating a Corporate Semantic Web. Contemporary business is based on huge amount of information and extracting right information at right time is a difficult and tedious task. By applying semantics within structured (ERP, Billing, Financial, HR systems) and unstructured data (email, fax, office documents) we can take business decision on the basis of overall organization knowledge base. [13].

2.3 A Focus Group Approach

As Competitive Intelligence encompass perspectives from different professionals, a focus group approach is more interesting than one-to-one interviews because it captures the dynamics of group interaction and exploits this in an attempting to understand a topic. [9]. A focus group is a way of collecting qualitative data, which involves engaging a number of people in informal group discussions around a particular topic or set of issues [10]. This qualitative research approach is effective in determining peoples views, feelings and opinions [24].

The main advantage of focus group interviews lies in the informal nature of the method, where instead of asking questions to each participant, the moderators encourage interaction between group members[10]. Because of these characteristics, a focus group is directly related to Socio-Technical Design, which argues that whenever a new system is being designed equal weight should be given to social and technical factors. [19].

3 CONTRIBUTION

This study presents how a focus group was used to raise requirements for a competitive intelligence software which should be based on open data and focused on mining companies. As geological formation does not depend on countries or borders, as soon as potential areas for mineral exploration are discovered, mining companies need to create multiple scenarios to support the decision process before making investments. Mineral exploration comprises extraction, beneficiation and pelletizing processes. Auxiliary processes include environmental management, industrial maintenance and logistics. Because of this characteristics, competitive intelligence in mining companies depends on professionals with different academic background. This would hardens the creation of a single survey able to get insights from each professional individually. In order to understand the set of information used by competitive intelligence professionals in a mining company, we performed a focus group as the data obtained is possibly richer and deeper than those stemming from individual interviews.

3.1 Focus Group Design

Script and Questions Before the meeting, we generated the focus group script to help the facilitator and the moderator during the meeting. Questions were selected to achieve the general purpose of focus group. Table 1 present questions asked in the focus group. We started with an opening question to create a welcoming and comfortable environment for the participants (question #1). Next, introductory questions were used to bring the participants closer to the main topic of the focus group (questions #2 to #3). A transition question was used to stimulate participants' memories with the intention to focus their attention on the key questions which were asked next (question #4). The key questions could lead directly to the focus group main question (questions #5 to #9). Finally, we used closing questions to capture any thoughts which were not mentioned before and conclude the session (questions #10 to #12).

Selection of Participants Participants - all from the same mining company - were selected based on their experience with competitive intelligence. Basically, three types of professionals were selected: strategy specialists, engineers and health, safety & sustainability professionals. We recruited a purposeful sample from personnel who had more than five years of experience with competitive intelligence. Invitation letters were sent to 11 (eleven) professionals and 07 (seven) of them attended the meeting. The size of our focus group adheres to the common norm of four (minimum) to twelve (maximum) participants per group [10]. Figure 1a shows the classification of participants by age and years of professional experience. Figure 1b shows the classification of participants regarding their specialties.

3.2 Focus Group Execution

In order to achieve the focus group purpose, we scheduled a meeting of two hours following the office hours. This way, participants had time to expose their thoughts without haste. At the beginning of focus group session, we informed participants about the study and gave them a written information form, including information on the aim of the study, anonymity issues, and a field for signing informed consent. During the session, a moderator facilitated the interview, while an assistant moderator was responsible for the audio-recording and noted the order in which participants spoke. As suggested by [11], the ending of focus group session occurred when all discussion triggers had been pursued and the participants had no further thoughts on the particular topics.

4 RESULTS

4.1 Major Difficulties

Participants reported difficulties to obtain information from countries which adopt non-Latin alphabets, such as: Cyrillic and Arabic Alphabets, Ideograms,

| Question #1	Have you ever been involved in environmental scanning analysis?
Question #2	What facts come to your mind when you think about environmental scanning analysis?
Question #3	Which sources have you ever used to obtain information for environmental scanning analysis?
Question #4	Suppose you were asked to conduct a survey using data from all external environment analyses you took part in. Which kind of information do you believe you would find in most of these analyses?
Question #5	If you had to analyze the external environment, what would be the most relevant information considering the politics and government of a region?
Question #6	If you had to analyze the external environment, what would be the most relevant information considering the social and economic indicators of a region?
Question #7	If you had to analyze the external environment, what would be the most relevant information considering the infrastructure and energy of a region?
Question #8	If you had to analyze the external environment, what would be the most relevant information considering competitors and suppliers of a region?
Question #9	If you had to analyze the external environment, what would be the most relevant information considering the community and local risks of a region?
Question #10	Suppose you were asked to conduct a survey using data from all environment analyses you took part in. Please reflect on the information that could have affected decisions if they had been present in the analyses that you performed.
Question #11	If the CEO of your company asked for your opinion on crucial information for understanding the external environment, what would you say?
Question #12	Would anyone like add anything to our discussion?

Table 1: Focus Group Questions

among others. Participants also mentioned difficulties to obtain data from competitors, such as: world presence, products specification, marketing analysis, etc. Another trouble reported by participants is when they try to obtain historical demand for products. Finally, trying to obtain a geo-referenced list of sacred places inside the country was mentioned as another barrier by participants.

4.2 Data Sources

Participants mentioned the existence of specific and generic datasources. Specific Datasources may vary depending on the country, business or competitor being analyzed. Generic Datasources can be used in multiple analysis, independently of country, business or competitor. Participants also separated datasources applicable to multiple segments from those applicable uniquely to their segment.
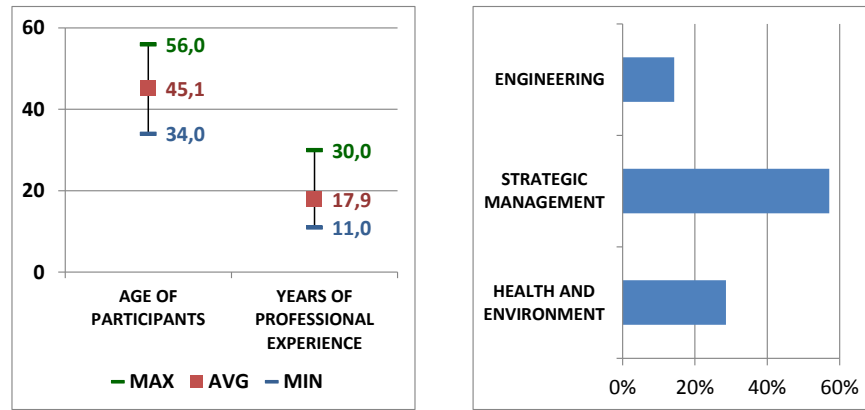
(a) Age and Experience of Participants

(b) Specialties of Participants

Fig. 1: Focus Group Demographics
Source: Author

Figure 2 presents a matrix organizing datasources by category and applicability. This figure indicates that results of this focus group can be used to formulate and design valid large-scale studies that will close relevant knowledge gaps for industries in segments other than mining.

4.3 Variables

The focus group identified sixty-seven variables as relevant to competitive intelligence process in mining industries. We based on [20] to classify variables according to their data source type. However, in order to suite the variables identified during the focus group, we needed to add three new categories. Figure 3 presents the complete list of information grouped by this new list.

5 DISCUSSION

Taking the objective of social-technical design into account, human needs must not be forgotten when technical systems are introduced [18]. Thus, a competitive intelligence software to mining industry should address the items below.

Firstly, users should be able to choose the relevant variables for each analysis and, from this choice, the system should visually indicate the countries whose variables are aligned to the established criteria. In mining industry, each business evaluation carried out by competitive intelligence professionals may explore different aspects. For specific reasons of business, professionals may perform analyzes focusing on structural issues in a given country. In this scenario, infrastructure's variables may be more relevant. In a different scenario, the focus may

Applicable to

Any Segment   Metals Segment

Category

Generic

GOOGLE; DOING BUSINESS PROJECT; UNITED NATIONS; WORLD HEALTH ORGANIZATION; WORLD TRADE ORGANISATION; CENTRAL INTELLIGENCE AGENCY; DELLOITE; REUTERS; INTERNATIONAL MONETARY FUND; EDGARD; SEDAR; ABNT; SCIENCEDIRECT; SCOPUS; U.S. CENSUS BUREAU; NASA; GREENPEACE; FRASER INSTITUTE; WIKIPEDIA

LONDON METAL EXCHANGE; CRU; SNL; BROOKHUNT; U.S. GEOLOGICAL SURVEY; STEEL BUSINESS BRIEFING; ESRI

Specific

GOVERNMENTAL DEPARTMENTS; SOCIAL WELFARE DEPARTMENT; CENTRAL BANK; TREASURY DEPARTMENT; STOCK EXCHANGE; RESEARCH INSTITUTES; INSTITUTIONS WEBSITES; COUNTRY'S LEGAL ADVISORY; DIPLOMATIC SERVICES; PRICE QUOTATION; BROKERS ANALISYS; CULTURAL ASPECTS; COMPETITORS; NGO'S; PATENT OFFICES

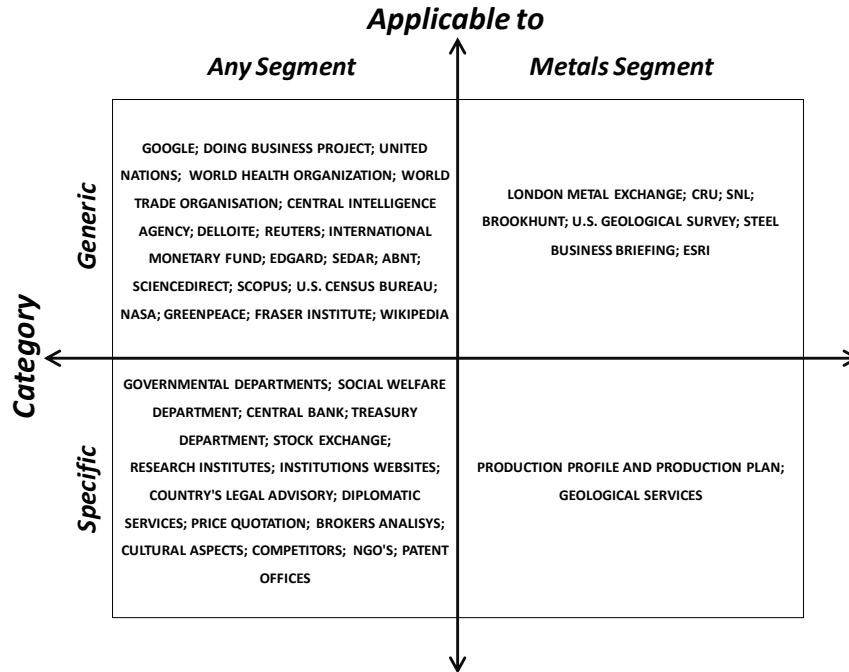PRODUCTION PROFILE AND PRODUCTION PLAN; GEOLOGICAL SERVICES

Fig. 2: Matrix of Datasources

be on demographic issues such as education level or human development index (HDI) of a specific country.

Secondly, users should be able to determine scores to each collected variable. It was mentioned in the focus group - and there was consensus between participants - that credibility of variables may vary according to data source. We confirmed that a single variable can come from different data sources with different values.

Finally, open data can be used to support users. From the variables identified during focus group, about 79% were classified as Open Data. We based on guidelines proposed by [2] to classify variables as open data. We created a dataset including the variables mentioned in the focus group and made this dataset available in RDF and acessible via SPARQL queries.

6   RELATED WORK

We found two groups of studies whose findings are related to this work. First group of studies [1, 15] is related because corroborate the level of importance of the information identified by competitive intelligence professionals in this study.

Demographic Data	- AIDS INDEX - CORRUPTION INDEX - COUNTRY GOVERNANCE REGIME - COUNTRY RELIGION - COUNTRY ECONOMIC SEGMENTS - COUNTRY SIZE - ETHNICITY - FORECASTED POPULATION - HDI	- HOSPITAL BEDS - LITERACY BY AGE - LITERACY INDEX - MINISTRIES - POPULATION - SOCIAL ORGANIZATION - UNEMPLOYMENT INDEX - WAR HISTORY
Macro Economic Data	- COUNTRY RISK LEVEL - COUNTRY'S MINERAL PRODUCTION - EDUCATION EXPENDITURE - EXCHANGE - FORECASTED DEMANDA - FORECASTED GDP - FORECASTED PRICES - FORECASTED SUPPLY	- GDP - HEALTH EXPENDITURE - INDUSTRIALIZATION LEVEL - INITIAL COSTS - PER CAPTA GDP - RATINGS
Data About Competitors	- COMPETITORS FINANCIAL PERFORMANCE - COMPETITORS GEO-REFERENCED LOCATION - COMPETITORS TOTAL INVESTMENT - COMPETITORS INVESTMENT PLACES - COMPETITORS PRODUCTS SPECIFICATIONS - TECHNOLOGY USED BY COMPETITORS	
Stocks and Shares	- COUNTRY'S MAIN ECONOMIC GROUPS - IMPORT AND EXPORT TRADES	
Internal Data	- CAPEX AND OPEX - CUSTOMERS DISTANCE - SUPPLIERS DISTANCE	
Geo-Referenced Data	- AIRPORTS - HOTELS - MINERAL RESOURCES - NEIGHBOURS COUNTRIES - PORTS	- RAILWAYS - SACRED PLACES GEO-REFERENCED LOCATION - TOPOGRAFY
Legal Issues	- ENVIRONMENTAL LAW - HEALTH LAW - LEGAL OBBLIGATIONS	- MINERAL LEGISLATIONS - TAXES - TECHNOLOGY OR PROCESS PATENTS
Infrastructure Data	- ARABLE LAND - EARTHQUAKE RISK - ENERGY DISTRIBUTION GRID - ENERGY MATRIX - FLOODS RISK	- HURRICANES RISK - POPULATION WITH ACCESS TO WATER - TSUNAMIS RISK - VOLCANOES RISK - WAR RISK

Fig. 3: Information List by Data Type

Impacts of the mining industry (economic, political, social and environmental) are described and the responses to these impacts discussed. Second group of studies [12, 26] is related because present a method to predict the capacity mining. Innovation, economic level, mineral resources production and the degree of coordination between industry and environment are shown as variables influencing mining industry in the future.

7 CONCLUSION

This paper presented both the information and the datasources used by competitive intelligence professionals in mining industries. Information required by those professionals can be classified in eight categories: demographic data; macroeconomic data; data about competitors; stocks and shares; internal data; spatial data; legal issues and infrastructure data.

In the light of social-technical design, a focus group helps the creation and development of knowledge. Thus, it can be considered as an appropriate instrument for this kind of study because while it makes it possible to generate outcomes combining perspectives from people of different academic background. Moreover, it also captures the dynamics of group interaction.

We concluded that it is possible to use Open Data technologies to gather and integrate information for competitive intelligence. However, governments and institutions are still encouraged to release their information using open data patterns in order to increase the percentage of information shared globally.

Although we believe our findings should have a certain transferability potential to competitive intelligence in segments other than mining industry, studying the relationship between organization and environment was never the purpose of this study.

References

[1] Banks, G.: Mining multinationals and developing countries: theory and practice in papua new guinea. Applied Geography 13(4), 313 – 327 (1993), `http://www.sciencedirect.com/science/article/pii/014362289390035Y`

[2] Berners-Lee, T.: Linked data. `http://www.w3.org/DesignIssues/LinkedData.html` (2009), accessed: 2014-03-10

[3] Bouthillier, F., Shearer, K.: Assessing Competitive Intelligence Software: A Guide to Evaluating CI Technology. Information Today (2003)

[4] Chen, H., Chau, M., Zeng, D.: {CI} spider: a tool for competitive intelligence on the web. Decision Support Systems 34(1), 1 – 17 (2002)

[5] Chen, H., Wu, Z.: On case-based knowledge sharing in semantic web. In: ICTAI. pp. 200–207. IEEE Computer Society (2003)

[6] Cronin, B., Overfelt, K., Fouchereaux, K., Manzvanzvike, T., Cha, M., Sona, E.: The internet and competitive intelligence: A survey of current practice. International Journal of Information Management 14(3), 204 – 222 (1994)

[7] Drott, M.: Personal knowledge, corporate information: The challenges for competitive intelligence. Business Horizons 44(2), 31 – 37 (2001)

[8] Eisenberg, V., Kanza, Y.: Ruby on semantic web. In: Abiteboul, S., Bhm, K., Koch, C., Tan, K.L. (eds.) ICDE. pp. 1324–1327. IEEE Computer Society (2011)

[9] Hyde, A., Howlett, E., Brady, D., Drennan, J.: The focus group method: Insights from focus group interviews on sexual health with adolescents. Social Science and Medicine 61(12), 2588 – 2599 (2005)

[10] Hyland, S., Haugen, A.S., yvind Thomassen: Perceptions of time spent on safety tasks in surgical operations: A focus group study. Safety Science 70(0), 70 – 79 (2014)

[11] Hyland, S., Haugen, A.S., yvind Thomassen: Perceptions of time spent on safety tasks in surgical operations: A focus group study. Safety Science 70(0), 70 − 79 (2014)

[12] Jingjing, Y., Rongge, X., JingHua, S.: Dynamic assessment on sustainable development capacity of mining cities by bp neural network. In: Information Engineering and Computer Science, 2009. ICIECS 2009. International Conference on. pp. 1–4 (Dec 2009)

[13] Khan, A., Hussain, M.: Application of semantic web in e-business and telecommunication. In: Advanced Computer Control, 2009. ICACC '09. International Conference on. pp. 513–517 (Jan 2009)

[14] Mariadoss, B.J., Milewicz, C., Lee, S., Sahaym, A.: Salesperson competitive intelligence and performance: The role of product knowledge and sales force automation usage. Industrial Marketing Management 43(1), 136 − 145 (2014), special Issue on Integrating marketing and operations for business sustainability

[15] Mason, C.M., Paxton, G., Parsons, R., Parr, J.M., Moffat, K.: for the benefit of australians: Exploring national expectations of the mining industry. Resources Policy 41(0), 1 − 8 (2014)

[16] McCrohan, K.F.: Competitive intelligence: Preparing for the information war. Long Range Planning 31(4), 586 − 593 (1998)

[17] Moresi, E.: Memoria Organizacional e Gestao do Conhecimento. In: Tarapanoff, K. (ed.) Inteligencia, Informacao e Conhecimento, pp. 277–301. IBICT (2006)

[18] Mumford, Enid: The story of socio-technical design: reflections on its successes, failures and potential. Information Systems Journal 16(4), 317–342 (Oct 2006), `http://dx.doi.org/10.1111/j.1365-2575.2006.00221.x`

[19] Mumford, E.: New treatments or old remedies: is business process reengineering really socio-technical design? The Journal of Strategic Information Systems 3(4), 313 − 326 (1994), `http://www.sciencedirect.com/science/article/pii/0963868794900361`

[20] Nettleton, D.: Chapter 3 - incorporating various sources of data and information. In: Nettleton, D. (ed.) Commercial Data Mining, pp. 17 − 47. Morgan Kaufmann, Boston (2014)

[21] Passant, A., Laublet, P., Breslin, J.G., Decker, S.: Semslates: Improving enterprise 2.0 information systems using semantic web technologies. In: CollaborateCom. pp. 1–10. IEEE (2009)

[22] Pinto, V., Parreiras, F.: Enterprise linked data: A systematic mapping study. In: 11th International Workshop on Web Information Systems Modeling (WISM 2014). Springer, Atlanta, USA (2014)

[23] Queyras, J., Quoniam, L.: Inteligencia Competitiva. In: Tarapanoff, K. (ed.) Inteligencia, Informacao e Conhecimento, pp. 73–98. IBICT (2006)

[24] Simigiu, A.: Investigating the views of civil society on early pregnancy through the focus group method. Procedia - Social and Behavioral Sciences 127(0), 219 − 224 (2014)

[25] Tarapanoff, K.: Informacao, Conhecimento e Inteligencia em Corporacoes: relacoes e complementaridade. In: Tarapanoff, K. (ed.) Inteligencia, Informacao e Conhecimento, pp. 19–36. IBICT (2006)

[26] Townson, P., Wrigley, C., Matthews, J., Bucolo, S.: Making sense of purpose, direction and innovation: An embedded design led innovation case study in the australian mining industry. In: Design Management Symposium (TIDMS), 2013 IEEE Tsinghua International. pp. 268–276 (Dec 2013)

[27] Willer, M., Dunsire, G.: Bibliographic Information Organization In The Semantic Web. Elsevier (2013)

[28] Xu, J., Zhu, Q., Li, J.Z., Zhang, P., Wang, K.: Modeling and implementation of unified semantic web platform. In: Web Intelligence. pp. 603–606. IEEE Computer Society (2004)