Information Design for "Weak Signal" detection and processing in Economic Intelligence: A case study on Health resources

Sahbi Sidhom, Philippe Lambert

Available for free online at https://ojs.hh.se/

Journal of Intelligence Studies in Business 1 (2011) 40-48

Information Design for "Weak Signal" detection and processing in Economic Intelligence: A case study on Health resources

Sahbi Sidhom * and Philippe Lambert * *

* Loria/Kiwi & Nancy Université, 4 Rue Ravinelle, 54000 Nancy, France
sahbi.sidhom@loria.fr

** VinaLor, Nancy Université, 4 Rue Ravinelle, 54000 Nancy, France
philippe.lambert@vinalor.fr

Received 20 February 2011; received in revised form 22 November 2011; accepted 25 December 2011

Abstract: The topics of this research cover all phases of "Information Design" applied to detect and profit from weak signals in economic intelligence (EI) or business intelligence (BI). The field of the information design (ID) applies to the process of translating complex, unorganized or unstructured data into valuable and meaningful information. ID practice requires an interdisciplinary approach, which combines skills in graphic design (writing, analysis processing and editing), human performances technology and human factors. Applied in the context of information system, it allows end-users to easily detect implicit topics known as "weak signals" (WS). In our approach to implement the ID, the processes cover the development of a knowledge management (KM) process in the context of EI. A case study concerning information monitoring health resources is presented using ID processes to outline weak signals. Both French and American bibliographic databases were applied to make the connection to multilingual concepts in the health watch process.

Keyword: Economic Intelligence, Business Intelligence, Information Design, Weak Signals

1. Introduction

On November 26th. (2010), the University of California officially launched its laboratory project on "Information Design" (ID). The project aims to develop knowledge exchange between different actors through applications for new media platforms such as Ipads in networks or Iphone Technology. Beyond the innovative aspect of this project, we note that the ID is the projection of an important "*Prospective Approach*" in the Anglo-Saxon research world. This point is reinforced by the comparison of the scientific literature on the issue. Since the 70s, research teams have specialized on the connections' between the graphical representation of information and its interpretation. One of the representation techniques that have been developed is "*spatial*" information across neuron networks. Especially in France, this approach has been somewhat vulgarized at first as in the example of mind-maps and Mind Mapping in education research. In recent years, this research focus has been applied to data mining using data from the Web (Web Mining). It helps to develop new knowledge from large text themes. This technique is increasingly interested in leaders who have the responsibility to detect topics that can have been missed in a linear reading. In the field of EI studies, the implicit properties on analysis take on the name "weak signals" (WS) (the explicit properties are "strong signals").The detection of WS allows the user to take better account of the environment in a dynamic sense and to build foresight ("To prepare today for tomorrow"). However, the connection between ID and WS detection requires the development of a complex methodological process. This is the topic of this paper. The first part of the paper defines the meaning of weak signals and processes through a strategic approach. The second part presents the logic in ID processes that tends to present varied graph data sets while facilitating the appropriation of "semantic" properties. The last part is matched to a case study on "strategic" health watch for which we use mapping and the visualizing of information. The study was able to detect a number of weak signals on scientific and technical information with the assistance of ID.

2. Mapping information for detecting weak signals (WS)

Anticipating strategic failures is one of the most common issues in EI studies. Market volatility, uncertainties about property prices and economic change are signals that announce future crisis and breaks from crisis. These breaks may be opportunities or threats in a changing world of economics where the faculty of anticipation becomes a powerful strategic advantage for companies. In 1970, Ansoff discussed the concept of WS in his first paper on the subject, entitled "*Managing Strategic Surprise by Response to Weak Signals*" (Ansoff, 1975). He considers the WS as corollary of organizational factors in the company, especially due to environmental turbulence as compared to the formulation of corporate strategy. In a following paper he specified the nature of WS, by defining it as "a warning (external or internal), events and developments that are still too incomplete to allow for an accurate estimate of their impact and/or to determine a full adapted response" (Ansoff, 1985). In what follows, we set out to determine the theoretical framework and application of WS.

2.1 Theoretical framework: weak signals (WS)

Any company can engage in a strategic process.
The specificity of the WS lies also in its potentiality. If one considers the famous *S-curve, which* describes the four phases of the product's life (birth, growth, maturity and decline), we can imagine that WS is a precursor of a new trend upstream of the cycle. Hence, the importance of WS detection in a logic of competitiveness. The term "*signal*" is ambiguous. If one refers to the definition in the Treasury of the French language (ATILF) for the word "signal" we find: [in French] "Signe convenu par lequel quelq'un Donne une information un avertissement á quelqu'un le moment de faire quelque chose". Ansoff's: [in English]: A sign by "*proactive*" value: to capture WS by the decision-maker via the channel of intuition (i.e. spontaneous knowledge of the environment) to cause a request for additional information (i.e. explicit formulations) from these signals. Another contribution to the question of weak signals was made by Coffman who has worked on various aspects of the problem. For him, a WS is defined as (Coffman, 1997): An idea that affects the way we trade and the environment in which we work; a novelty and a surprise in terms of receiving signals, a noise and other signals, sometimes difficult to detect among noise and other signals, an opportunity or a threat to the organization, often made fun of by the "knowledge holders" or experts, weak signal with a substantial period of time before it matures and becomes a strong signal, therefore, this signal represents an opportunity to learn, grow and evolve. Coffman (1997) also said that the WS could be of three types: *supra-perceptual signal, perceptible signal* but not recognized by our mental models, and *recognized signal* by our mental models and by which our change in behavior. In France, H. Lesca (2001) proposes a list of characteristics that define a WS, which

is close to that of Ansoff. A signal can be classified as WS if it is fragmentary, embedded in a mass of useless information (or noise), an apparent weak and ambiguous meanings, could not be seen, an apparent low usability, and low "palpability". In synthesis of these definitions and presentations, we can consider that a weak signal is characterized by: a temporal discontinuity of its discovery, but also by the fact that it causes a shift (or breaking) in the facts found by the receiver to arouse/create measurable interest in the future. The researchers found, in the notion of breaking or the "discontinuity", the reason for the information flow and design that someone provides information, a warning to someone; someone tells it's the time to do something. Precisely the opposite that constitutes the "weak signal" in a strategic watch process. The transmitter of the information detected as WS do not expect the risk that competitors become aware of the potentially innovative nature of the information given. The adjective term "weak" is also a sematic problem. The "weakness" of the signal is opposite to the potential of information designated by this term. The term "weak signal" is defined by "a high potential for new innovations". Due to this we propose to transform the term "sign" above as in "weak signal" to "*Latent Warning Sign*" (LWS) for designated information in a strategic context.

2.2 Application framework: Knowledge discovery

For nearly a decade, several teams of researchers across the Atlantic have focused on the subject of Information Design (ID). The concept, however, remained forgotten in France until recently. Visualizing information has great potential advantages. Eppler and Burkhard (2007) gives six main reasons for why it is important to give priority to this area: it motivates the receiver, presents new perspectives, develops memory, encourages the learning process, captures the attention of the receiver, and allows structuring and coordinating of communication. Many definitions present ID as an art, the art to direct information to create meaning. Graphic productions goes together with significant creativity, with formatted, colorful, animated and multiform information. In addition to the purely aesthetic side of this approach, ID contains intrinsically a new way of thinking about information and could be summarized as Karabeg (2002) did in "a new approach to information": He explains what ID is by proposing the image of a bus equipped with "candle flags" (Atilf, 2010). The bus represents a "modern culture" while the candles symbolize "traditional" information. We observe what the author means by this incongruence (dysfunction). However it can be surpassed by the use of ID. Moreover, "*modern culture*" is producing and consuming information on a large-scale. ID through the development of information technologies can act as a remedy to the problem of chronic "*infobesity*". On this epistemological ID, more technique is being added, proposed by men like Jotham Fry (2004) in his thesis entitled "*Computational Information Design*". The issue of work is to propose a methodology for data visualization and offer a comprehensive set of graphical representations to give sense to implicit relations between connective data. Jotham Fry (2004) presents a classic seven step process for ID to ensure the transition from data to knowledge, *acquire*: acquisition of data from any medium, *parse* or split; cutting to provide a structure of the data and order, *filter*: filtering to select only relevant data, *mine*: the search where you place the data into a mathematical context, *show*: representation where it is determined, that a simple representation of data can take, *refine*: refining to change the simple representation to more and advanced visual renderings, and *interact*: interaction by adding methods for manipulating data through visualization.

Besides the purely aesthetic side of information (or infographic), ID is at the crossroads of several fields of scientific applications. This includes the fields of visualization techniques to computer graphics for greater knowledge. We also see the impact ID has in psychology and semiotics. This development is to perfect the cognitive and physiological theories of visual perception and cultural factors that come into account in the process of information visualization. Ultimately, ID enables the end-user, usually an expert whose skills enable him to interpret the data represented as graphs, to generate links between data and knowledge. This knowledge discovery is not the ultimate goal of the logic of ID. On the contrary, the new application aims to refocus the attention of the user to historical data previously unnoticed. As a result, a new watch cycle will begin on data previously "unnoticed". In this case study, the aim is to illustrate the relation of ID and Knowledge Data Discovery (KDD). The "Latent Warning Sign" (LWS) is here a key component of this new application by the emergence of thematic relations that may improve the strategic watch process.

3. The Case study: the health heterogeneous resources – project « cronisanté »

During the conference SIIE'2010, there was a study about reflections on "chronic diseases

management (project 2007)", by the High Council of Public Health (HCSP) with INIST-CNRS in France, to establish an "Information System for Decision Support" powered by a strategic watch process on the health resources (Lambert & Sidhom, 2010). The HCSP was trying to identify how European health systems manage the problem of "chronic diseases". The approach to the problem is based on the WISP model (i.e. Watcher Information and Search Problem) developed by P. Kislin (2007). This model is the extension of a watch approach to describe the information needs and help the user (decision-maker and user) to formulate needs. In the context of this work, the formulation of needs has been directed towards the bibliographic references obtained after consulting a business database (*cf.* III.D). The strategic issue of this work, for the user, is to formalize the declarative rules, by: <<ISSUE>>: = if we do NOT act on the <OBJECT> and Knowing the state of the <SIGNAL>, then the risk is the expected <HYPOTHESIS>. Where: ISSUE is defined by an OBJECT of the environment, on which it is possible to act. A SIGNAL prompts the decision-maker to trigger the problem, an HYPOTHESIS, which is the risk, as, expected consequences, if left unchecked. The approach in view of the application is to better target the information needs of the project sponsor, the HCSP. To achieve this we translate the strategic issue in a series of dimensions related to the problem with a set of indicators on the Information Retrieval (IR) process. One can easily imagine, given the multidisciplinary nature of the working group and the specific interests of each expert, that the heterogeneity of the subject fields represent a problem in the collected information and for analysis.

3.1 Health resources: Semantic heterogeneity

"Too much information kills information" has become the favorite expression of responsible users at the time of information overflow. The ability to quickly extract relevant information, while providing added value, creates more robustness to any surveillance process. The concept of added value will here be understood as the annotation process to facilitate access to relevant information for the user (Sidhom, 2008). Indexing and reindexing by users (i.e. social tags, folksonomies, etc.) are in the list of tools for this process. Furthermore, the quantity of information, the heterogeneity of resources and information itself are a problem well known to designers of information systems. Schematically, we speak of a dual heterogeneity that is both semantic and syntactic. Example of this is the syntax for the heterogeneity of data storage formats (pdf, doc, xml, etc.), query languages and more generally across protocol data structure. The semantic heterogeneity is the differences between the interpretations of the real world inducing several terminology uses for the same reality (ontology, synonymy, etc.) (Goh, Bressan, Madnick and Siegel, 1999). Later, we will return to this problem by relying on examples from the bibliographic databases. The added value gained from the system use two aspects. The first is the addition of keywords or comments (i.e. social tags) by the user to information resources. This allows customization of information regarding documentary resources. These annotations can feed up the index in a system to improve the return rate for IR. The second aspect concerns the "Information Design" process (Ansoff, 1975). Several studies in the medical sector have shown that visual information influenced the decision-making both in a strategic situation (i.e. care policies) and a therapeutic condition (i.e. alternatives to hospitalization) (Elting, Martin, Cantor and Rubenstein, 1999), (Wyatt. 1999). This logic has not only to refine the conceptual goal in the system but also to support the iterative process: – information needs – IR – new conceptual indicators (proposing). This process can come from techniques of Knowledge Management (KM) and Mind Mapping.

3.2 From NLP to Information Visualization

Sidhom (2002) has developed a Morpho-syntactic Analysis Platform for automatic indexing and information retrieval (SIMBAD). It is composed of an Indexing Kernel (i.e. indexing process) that uses the noun phrases (NP) as descriptor in NL structures (i.e. to extract concepts) in text documents (and opens to multimedia associated to text descriptions). We use the definition of a noun phrase (NP) as defined by Le Guern (1989) to place a lexicon word in the discourse of universe, de facto, this word is ejected in extensional logic, and gives to NP a repository status, as a reality segment associated with it. In our context, the NP appears to be the bearer of a "semantic load", which makes it relevant and a central element to bibliographic information analysis. Around this semantic we search guides for our analysis on the actual corpus. Thus, grammar of NP recognition has three logic levels:

> 1°/ *Intentional level* (or natural language properties), is represented by the level N. Words are considered free predicates or as simple (i.e. the noun properties) or as complex (i.e. the noun properties modified by other units:

adjectival units A' (i.e. A' A|Adv+A|A+Rel, etc.), expansional preposition EP (i.e. EP Prep+N', etc.), etc.

2°/ *Intermediate level* (or taking into account the universe of discourses), it is represented by the level N'. It is the transition from the intentional to the extensional levels. Words are considered free predicates with a set construction of closed predicates to denote objects in the world(i.e. N' N+SP|N+A'|…|N)

3°/ *Extensional level* (or the NP and its complexity), it is represented by the level N''. It is the close operation using a quantifier that selects a specific element in the class N of nominal.

These are the existing objects in the world, referred objects or mind-constructed objects. In this work, the morpho-syntactic grammar of the NP has been rewritten for NooJ in two levels: firstly, the work was to reformat linguistic resources (dictionaries and grammars) resources in our possession. During a second time we developed the finite state transducer of the noun phrase. Labels existing dictionaries have been harmonized to match the syntactic graph of NP (FIG. 1-2).

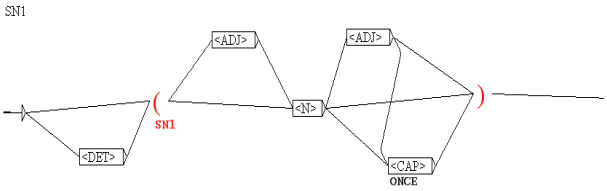

Figure 1: Syntactic graph of simple np in Nooj

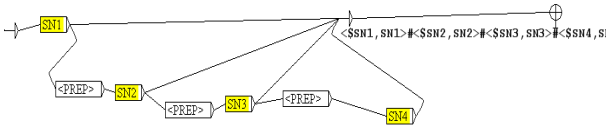

Figure 2: Embedded syntactic graph of complex np in Nooj

The graph provides numbered phrases identifying the fitting relations in syntagmatic level results (Lambert and Sidhom, 2010). In the logical use of the semantic concepts (i.e. NP and its properties) from the bibliographic records, the results on output graphs must be operated by an end-user. This gives the end user access to pure information, leaving him free to evolve in the concepts from a document process: visualizing information spaces fed by heterogeneous data sources. This is a support of an economic intelligence process and for a information design system (Lesca, Kriaa-Medhoffer and Casagrande, 2009), for the "*ChroniSanté*" project. In particular, in information surveillance activities, the process is a major vector for the emergence of significant associations after phases of collection, processing and analysis in a large mass of data and information. Several solutions to information mapping software are available. The tool we used is software under GNU General Public License (GPL3) called Gephi (http://gephi.org). It allows the visualization of complex networks.

3.3 Corpus study and indicator valorizations

As part of our core construction, we mainly searched bibliographic databases via the multi-application "Webspir", a tool that was replaced at the start of 2009 by the platform "OvidSP" (http://www.ovid.com) with features near equivalent but more robust for users. Three databases were selected for the constitution of bibliographic our entities: <Pascal[2]> <PsyInfo[3]> and <Medline[4]>. The choice to use these three sources on health information is justified by our aim to cover as fully as possible the thematic management on chronic diseases. The basic advantage of Pascal database is that it presents European references and includes records from the databases in public health (BDSP). The *Medline* database is centered on U.S. publications, such as *PsycInfo*, but with a broader theme in the social sciences. In synthesis and contrary to this logic which requires complete topics, our search equations were developed to deliver results to the widest possible extent: first, to cover all the sub-themes on "chronic disease" and, second, to identify new sub-themes which we had not originally thought of. The browsing on the three databases reported: 2097 references to *Pascal*, 6110 references to *Medline, and* 2177 references to *PsycInfo*. We subsequently refined our search to select only those published between 2001 and 2009 in French. The result consists of 397 references and 303 references in the duplication pass. These entities then will be the first synthesis of our work. The results of the IR process indicate that "*chronic disease*" is a new concept in France, because of the singularity of the model in the French health system. A second approach has motivated a second job on the database "Pubmed" as previously mentioned. The completion of the "*ChroniSanté*" project as a "decision support system" (DSS) or SIAD in French, was faced with a semantic problem: the rendering of the term "chronic disease" is as a concept purely Anglo-

Saxon, which brings a series of problems in a multilingual and "Translation Terminology". In fact, the completeness of the study involves a search process on multilingual literature to define the best concept and study consisting of what intersects. It is in this logic that the base "Pubmed" was viewed with a search for the term "chronic disease" in the title of the records. The result is 13,222 records. These had parallel entities in the English language with relation to the initial multi-base.

4. Connections to latent warning (or weak signals) in the id process

Applying automatic analysis (NooJ) on our data entities (ie. as the ID process in phases 1 acquire 2 parse), the complex graph of NP reported 1374 concepts (ie. as the ID phase 3 filter) including the smallest concepts (ie. the lemma N) to simple or complex concepts NP (ie. levels N' + N"). The advantage of this approach is to present to users the primary concepts (i.e. as the ID phase 4 mine) in information resources but also secondary concepts (i.e. the ID phase 5 representation) that the user does not necessarily think of in his research of indicators: the translation phase of a decision problem into an IR problem in the EI context. In this case, considering the concept of "patient" is the central tour theme. In practice, we tend to establish our search for indicators in a passive acceptation with concepts: "patient monitoring [(FR) *suivi du patient*]", "patient care [(FR) *prise en charge du patient*]", "patient education [(FR) *éducation du patient*]", etc. But not in an active acceptation, as "patient involvement [(FR) *implication du patient*]", "active participation of patient [(FR) *participation active du patient*]", etc. (i.e. as the ID phases 6 refine in iteration). The research of NPs in the titles of references highlighted ideas that apparently have no close relation with our themes, but which nevertheless appear several times in different references. In this case, on the theme of "cannabis consumption" it puts a link for "long term illnesses". Given these observations, we took for advantages of select matches, the longest in the NP. This corresponds to the fitting relation: the concept of fitting in (x y), x the longest and most informative NP; the concept fitted (zw), z the shortest and the least accurate NP. The richness of meaning that emerge, allows the identification of informational collection with relevant, complex and hierarchical concepts (NP). These characteristics may go unnoticed in the linear analysis of an entity. Thus, the process of the bibliographic entity records on the Platform NooJ showed satisfactory results based on the NP and its semantic properties (i.e. the relations of Tree (T: yx and xz, Fitting (F: xy w) and Belongs (B: n x / xy)) (Sidhom, 2002). Pascal: Produced by INIST-CNRS, PASCAL is an international and multidisciplinary database that identifies literature in Science, Technology and Medicine. PsycInfo is the Database of the American Psychological Association (APA) and provides access to journal articles (many are full text), book chapters and books, research reports and theses and dissertations in psychology and related fields (medicine, nursing, sociology, etc.), from the 19th century to today. Medline is a Bibliographic database produced by the National Library of Medicine (NLM-USA). It covers all biomedical fields: biochemistry, biology, clinical medicine, economics, ethics, dentistry, pharmacology, psychiatry, public health, toxicology, veterinary medicine.

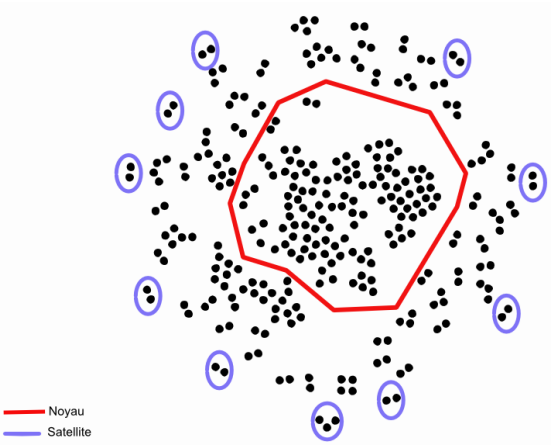

Figure 3: Semantic networks based on bibliographic records: nucleus and satellite connections.

Concerning the visualization of information (i.e. as the ID phase 7 interacts in iteration), we tested the application with Gephi Fruchterman-Rheingold algorithm (Card, Mackinlay and Shneiderman, 1999) on the results of extraction with NooJ. This algorithm of multi-scale force can calculate the force between two nodes and map complex networks. By its use, there is much emerging nucleus surrounded by satellite subsystems that can be considered non-central themes to the theme target. According to the analysis of information needs, the user can focus attention on these satellites nodes to be considered as "latent warning signs" themes (Weak signals in EI) and give them special attention (FIG. 3). By NooJ parsing, we present the graph results of the NP extraction in the French entity (FIG. 4).

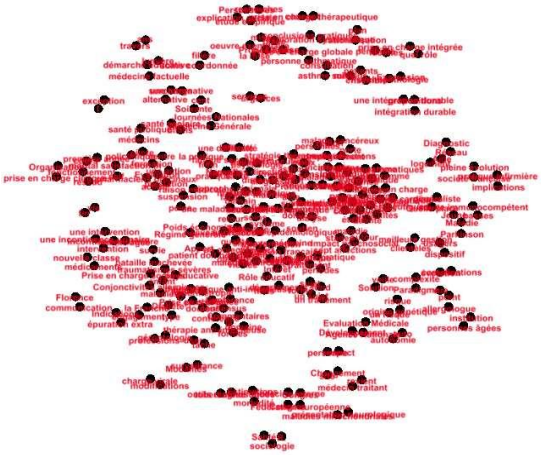

Figure 4: Visualization of nap semantic networks

Based on the analysis of the semantic network, we observe that the center of the graph (central nucleus) consists of terms related to the decision-making analysis: work that we completed in the process of EI. Thus, for the analyzed entity, it reprsents terms such as "coverage, care [FR: *prise en charge*]" or "chronic diseases [FR: maladies chroniques]", new terms like: ("hepatitis C", "cardiopathy", "asthma", etc.). We notice the relation that exists between nodes that represent the semantic connections between terms (FIG. 5). The usefulness of such a "visual structure" document for a user, an expert or a decision-maker, is considerable. It allows presenting an interactive document likely to bring new knowledge. In such a semantic logic, it allows to better understand the complex dimensions present when concerned with the surveillance or EI process.

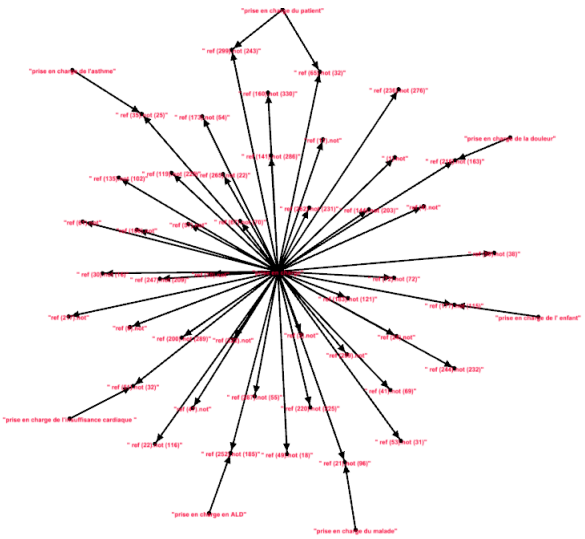

Figure 5: The np core (central nucleus) and themes.

The visualization of named entities in the data can also be positioned relative to the documentary logic. The graph makes it possible to show relations between concepts (NP) and document references (FIG. 6). Said differently, we can see which resources concentrated more concepts and which ones that is potentially relevant.

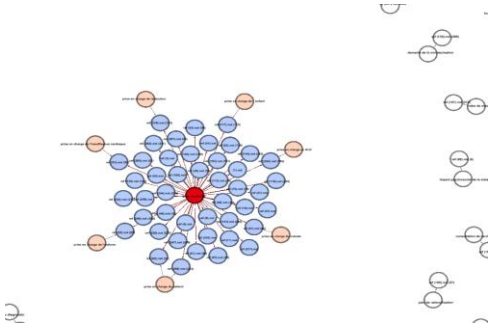

Figure 6: Connection between np and information resources.

The atomic structure for the nucleus concept, "coverage, care [FR: *prise en charge*]" is linked to the bibliographic references. We also note the secondary concepts like: "management of asthma [FR: *prise en charge de l'asthme]*", "chronic asthma [FR: *asthme chronique]*" and other concepts which appear in the peripheral area of the semantic network. This process was applied to the second data entity in English, with the same Logic: the extraction of NPs based on graphs modeling (Silberztein, 2005). Four hundred and twenty- nine mapped terms are returned. They are linked with their document resources. This approach requires of users a "proactive approach" through the research (mining) graphs proposed in order to detected new knowledge. In the logic presented here, it may refine the concepts from another cultural sphere. In application, the activation of a term (T) allows us to view notice records (N1. Ni) in connection with subjects (S1. Sm) and associated keywords (K1. Kj). The scenarios for identifying and refining may be multiple: as an example, from record (Ni) to keywords (K1. Kj) or vice-versa. The possibility to link the node "notice" to the source document by hyperlink feature allows the user to have access to the document environment in terms of interest with any subject.

5. Discussion and conclusion

The experience shows that the techniques we used require automation to achieve a state of performance, robustness and an acceptable level of efficiency. The "*Information Design*" (Ansoff, 1995) process defined as the "art and science of preparing information so it can be used by humans with efficiency and effectiveness." In the ID

process, we evaluated our study areas by the use of the following aspects: "graph(s)", "semantic network (s)", "project (s)" and "connection(s)", to translate them into clear, immediate and appropriate information for users. In our study, the user is often the observer, the analyst or expert and the decision-maker. For us, useful information is not the increase of information quantity, but on the contrary, the reduction of it by relevant information clusters to facilitate its reading and appropriation (Lesca, Kriia-Medhaffer and Casagrande, 2009). This has been discussed and treated throughout this paper explicitly as the application of ID processes in the context of our "chronic diseases" study (project "*ChroniSanté"*). For the "*Surveillance" process*, we found that the information visualized extracted from the concepts of NPs is useful to actors in a strategic project in numerous aspects. At first, information visualization facilitates document indexing and content of information systems (IS), information retrieval (IR) systems or decision support systems (DSS). As an example, for a bibliographic record or document to be analyzed, to extract noun phrases in the content we may convert them into tags. This solution allows any user of the document to present the key concepts in the information database (Lesca, Kriia-Medhaffer and Casagrande, 2009). This may again encourage a new Logic of "reindexing by users": The user-tags are automatically stored; the user will add subjective, objective and creative tags, to give added-value (Harboui, Ghenima and Sidhom, 2009). Second, the visualization of a semantic network (based on NPs concepts and properties) enables the production of new knowledge. Viewed nodes in a semantic network can indeed be analyzed in a working group or team to identify new topics related to business intelligence as convergence and divergence of represented subjects (i.e. decision-making needs). This is, to use the "Humbert Lesca" logic; heuristic processes allow a collective creation of meaning (Lesca, Kriia-Medhaffer and Casagrande, 2009). At third, in the surveillance, information and documentation processes, the visualization logic and the ID process can bring out potentially relevant data. On this point, it should refine results by a statistical analysis: the use of bibliometric indicators (Salton, Wong and Yang, 1975) as the TF-IDF (term frequency-inverse document frequency). For "*Economic Intelligence*", we see that the usefulness of the ID process goes beyond a simple and literal translation of IR indicators (i.e. the translation phase of a decision problem into an IR problem). For complex and "multilingual" semantic search, we can take the conceptual differences between terms, like: "chronic disease" or "chronic disease management" and "management of chronic diseases". Thus, we can show, by semantic visualizing of these concepts, the connections with the processed information (parsing, analysis and needs information). Also, we can establish multi-level intersections between information concepts and common or different "morphemes" to get shared meanings. Finally, on a technical level, the problem of heterogeneous information resources can be minimized with the addition of annotations and/or linguistic processing. NLP tools (as NooJ or others) will enable the processing of multi-format sources, and for non-textual documents (multimedia), analysis can be made on the annotations associated with the documents: developing a semantic homogenization, by counterbalancing the syntactic heterogeneity.

References

ATILF – CNRS. 2010. - Université Nancy 2, Le Trésor de la Langue française, http://atilf.atilf.fr/dendien/scripts/tlfiv5/advanced.exe?_8;s=66859215; [visIted date: Nov. 20, 2010].

H. I. Ansoff. 1975. "Managing strategic surprise by response to weak signals". *California Management Review*, vol. 18, no. 2, pp. 21–33.

H. I. Ansoff. 1985. *"Strategic response in turbulent environments"*. Handbook of business strategy.

S. Card, J. Mackinlay, and B. Shneiderman (Eds). 1999. *Readings in Information Visualization: Using Vision to Think*. S.F., CA: Morgan Kaufmann.

B. Coffman. 1997. Weak Signal Research, Part I: Introduction, *Journal of Transition Management*, MG Taylor Corporation, Jan 1, 2011, from http://www.mgtaylor.com/mgtaylor/jotm/winter97/wsrmatur.htm

B. Coffman. 1997. Weak Signal Research, Part II: Information Theory, *Journal of Transition Management*, MG Taylor Corporation, Retrieved Jan 1, 2011, from http://www.mgtaylor.com/mgtaylor/jotm/winter97/wsrmatur.htm

L. S. Elting, C. G. Martin, S. B. Cantor, and E. B. Rubenstein. 1999. "Influence of data display formats on physical investigators´ decisions to stop clinical trials formats on physician investigators´ decisions to stop clinical trials: prospective trial with repeated measures," *BMJ*, vol. 318, no. 7197, pp. 1527 -1531.

M. J. Eppler and R. A. Burkhard. 2007. "Visual representations in knowledge management: framework and cases". *Journal of Knowledge Management*, vol. 11, no. 4, pp. 112–122.

B. Jotham Fry. 2004. Computational information design. bfa Communication Design, in Computer Science Carnegie Mellon University. Massachusetts Institute of Technology, April, Retrieved 5 Jan from http://benfry.com/phd/dissertation-050312b-acrobat.pdf

C. H. Goh, S. Bressan, S. Madnick, and M. Siegel. 1999. "Context interchange: new features and formalisms for the intelligent integration of information," *ACM Trans. Inf. Syst.*, vol. 17, no. 3, pp. 270-293.

M. Le Guern. 1989. "Sur les relations entre terminologie et lexique," *Meta*, vol. 34, no. 3

A. Harbaoui, M. Ghenima, and S. Sidhom. 2009. "Enrichissement des contenus par la réindexation des usagers: UN état de l'art sur la problématique". In *Proceedings SIIE'09 -Systèmes d'Information ET Intelligence Economique, IHE Edition. Hammamet* (12-14 fév. 2009).

D. Karabeg. 2002. "Designing information design". *Information Design Journal*, vol. 11, no. 1, pp. 82–90, 2002.

P. Kislin. 2007. "Modélisation du problème informationnel du veilleur dans la démarche d'Intelligence Économique". *Thèse de doctorat*: université de Nancy2. 5 Nov.

P. Lambert, S. Sidhom. 2010. "Knowledge extraction and visualization: case study on ChroniSanté project in France ". In *Procedings SIIE'2010, Sousse (Feb. Tunisia). IHE Edition.*

H. Lesca. 2001. "Veille stratégique: passage de la notion de signal faible à la notion de Signe d'alerte précoce," in *Colloque VSST*, pp. 98–105.

H. Lesca, S. Kriaa-Medhaffer, and A. Casagrande. 2009. "Veille stratégique: Un Facteur d'échec paradoxal largement avéré: la surinformation causée par l'Internet. Cas concrets, retours d'expérience et piste de solutions". in *Proceedings SIIE'09 - Systèmes d'Information et Intelligence Economique, IHE Edition. Hammamet* (12-14 fév. 2009).

G. Salton, A. Wong, and C. S. Yang. 1975. "A vector space model for automatic indexing," Communications of the ACM, vol. 18, no. 11, pp. 613–620.

S. Sidhom. 2002. "Plate-forme d'analyse morpho-syntaxique pour l'indexation automatique ET la recherche d'information: de l'écrit vers la gestion desconnaissances". *Thèse de doctorat*, Université Claude Bernard Lyon1.

S. Sidhom. 2008. **"** Approche conceptuelle par un processus d'annotation pour la représentation et la valorisation de contenus informationnels en intelligence économique (IE)

contenus informationnels en intelligence économique (IE)," In *Proceedings SIIE'08 - Systèmes d'Information et Intelligence Economique, IHE Edition vol.1, Hammamet* (fév. 2009).

M. Silberztein. 2005. "NooJ: a linguistic annotation system for corpus processing," in *Proceedings of HLT/EMNLP on Interactive Demonstrations*, p. 11.

J. Wyatt. 1999. "Same information, different decisions: format counts," *BMJ*, vol. 318, no. 7197, pp. 1501 -1502.