



XEW 2.0: A Competitive Intelligence System for Big Data Analysis in Cloud

Amine El Haddadi^{1,3}, Anass El Haddadi², Abdelhadi Fennan³, Benard Dousset¹,
Zakaria Boulouard⁴
¹IRIT, Toulouse University – Toulouse, France
²ENSA, Mohammed First University – Al-Hoceima, Morocco
³FST, Abdelmalek Essaadi University – Tangier, Morocco
⁴FS, Ibn Zohr University – Agadir, Morocco

Abstract: Competitive Intelligence (CI) is the set of coordinated researches, treatments and distribution of useful information to stakeholder to-wards action and decision-making. In order to enable users to search, monitor, validate and rebroadcast strategic information, we provide our new tools XEW 2.0, which can be helpful for them in their executive travels. In this paper, we focus on the architecture, multilayer model and service of the CI system – XEW 2.0 to describe our approach to treat different data-sources (Patents, Paper, etc.) in cloud computing.

1 Introduction

Nowadays, companies are faced with external risk factors linked with an increased competition market place – we know that markets are extremely dynamic and unpredictable: new competitors, mergers and acquisition, sharp price cuts, rapid changes in consumption patterns and values, weak brands and their reputation.

CI is a discipline to better anticipate risks and identify opportunities. Fifteen years after the canonical definition in French proposed by Martre [1], CI is still a concept with unstable borders. The last few years have seen a multiple definition of CI: from definitions oriented mapping process practice of CI, strategic vision of the CI to others including the concepts of knowledge management, collective learning and cooperation [2].

In the context of our work, we retain the concept of CI as it was defined by the Society of Competitive Intelligence Professionals (SCIP): Competitive Intelligence: A systematic and ethical program for gathering, analyzing, and managing external information that can affect your company's plans, decisions, and operations. Put another way, CI is the process of enhancing marketplace competitiveness through a greater -- yet unequivocally ethical -- understanding of a firm's competitors and the competitive environment. Specifically, it is the legal collection and analysis of information regarding the capabilities, vulnerabilities, and intentions of business competitors, conducted by using information databases and other "open sources" and through ethical inquiry. Effective CI is a continuous process involving the legal and ethical collection of information, analysis that doesn't avoid unwelcome conclusions, and controlled dissemination of actionable intelligence to decision makers.

In CI process multivariate techniques are currently well controlled for all available quantitative data, on condition that DBM be suitable, the database schema be adopted and data be of the highest quality (homogeneous, current, complete...). It is always possible to extract the relevant data to a database custom built for multidimensional analysis. But, textual data from all electronic sources (the scientific databases, the patents databases, the press, RSS, Internet, intranet, forums...) is difficult to implement: data sources have different formats or are even unstructured, they are distributed, heterogeneous and encountered many particular / singular cases, particularly when we analysis a topic from different points of view (science, technology, news, etc...).

To standardize the multidimensional analysis text data from all sources, we propose an unified structure [3, 4] for storing all relationship items encountered in analysis documents. This method allows the construction of tridimensional analysis among variable and a time. Since 2001[5], a first tool Xplor has been suggested to upload this type of structure in client server, to perform a custom search tools through various graphic restitution of the results. All text data are then in the same structure and are therefore common tools of interactive investigation. An improved version of Xplor appeared in 2007[6] and in order to enable users to search, monitor, validate and rebroadcast strategic information, we provide our new mobile tools XEW 2.02.0, which can be helpful for them in their executive travels.

The rest of paper is structured as follows: first, we identify in section II the literature view of CI and our CI process. In section III, we will explain the architecture of XEW 2.0 in cloud computing. Section IV summarizes and assesses the approach.



2 Literature view

2.1 Competitive intelligence

CI is not an innovative practice created exclusively for the development of technology and information society. Each country has led a process that can be called CI.

A broad definition of CI is a set of coordinated actions of research, treatment and distribution of useful information to stakeholders to enable the action and decision making. More, CI is a legal business practice, as opposed to industrial espionage which is illegal. Generally, the focus is on the external business environment [7, 8].

The concept of CI is now in emerging phase, but nevertheless it is still necessary to produce a reference framework based on a multidisciplinary approach [9, 10]. In our research team, we coordinate the process of CI around three concepts: strategic analysis, environmental scanning and information system. Fig. 1 shows a hierarchy illustrating the coordination of these three concepts.

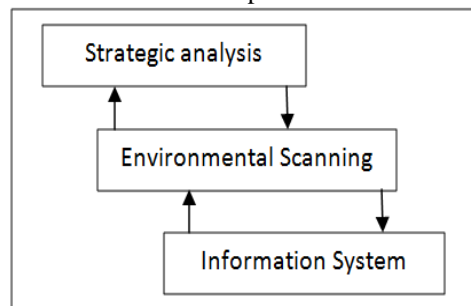


Figure 1. Coordination of concepts related to CI [11]

- Strategic analysis: define the informational needs of the company on its environment in order to reduce the uncertainty and to simplify decision making.
- The environmental scanning: will guide and steer the process of collecting, processing, analyzing and disseminating information to meet the expressed needs.
- The information system: is used to support the various activities from data collection to data analysis and reporting.

2.2 Competitive intelligence process

The Competitive Intelligence System (CIS) integrates processes of strategic analysis and environmental scanning based on information systems (cf. Figure 2): 1. Develop, 2. Choose, 3. Identify & Prioritize, 4. Identify & Select, 5. Collect & Evaluate, 6. Organize & Remember, 7. Validate & Stream and 8. Analyze & Interpret.

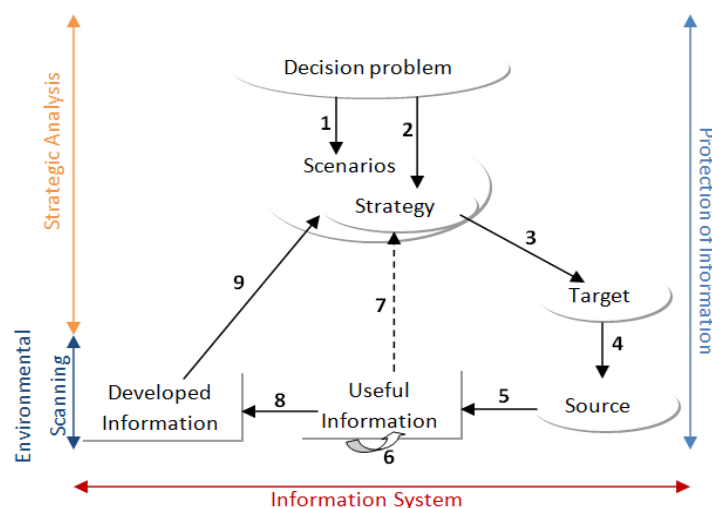


Figure 2. Competitive Intelligence process

The four essential steps of the intelligence cycle are included in this schematic presentation: understanding the need, researching and gathering information, processing information, and disseminating information.



3 XEW 2.0 system architecture

The CIS XEW 2.0 aims to improve decision-making in all aspects in business life, particularly offensive and innovative decisions. XEW 2.0 based on a multidimensional analysis model [11], whose objective analyzed the information environment in all dimensions of a decision problem, with the exploitation of information by analyzing the evolution of their interactions. Our approach combines two methods: knowledge discovering in text (KDT) and environmental scanning.

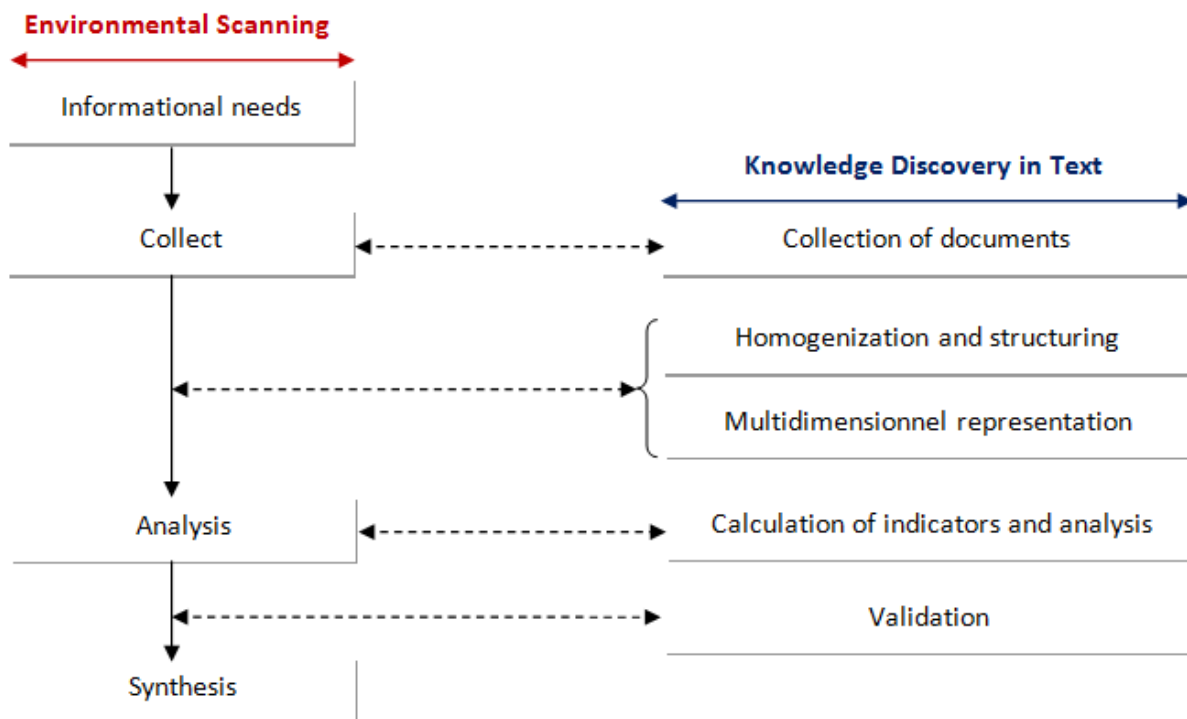


Figure 3. Coupling of Environmental scanning and KDT [11]

The dynamic aspect is vital to any analysis in the context of CI. These dynamics include continuous monitoring of the business environment in order to detect its changes and developments. The proposed information system, based on an exploratory multivariate analysis model: “the relational aspect and the time dimension”. It is based on extracting knowledge from textual data by analyzing relational data and their evolution. This model allows time specification to situate events, strategy and actions as well as in: the past by reconstructing the chronology; the present-oriented time to detect weak signal and the future to detect relationships in network, such as partnerships, alliances, mergers, acquisitions, co-citations, co-signatures, co-occurrences of all kinds.

3.1 System architecture

The architecture of our platform consists in 4 main services: **Sourcing Service**, which is to collect information (scientific database, Patents database, RSS, Blogs, etc.). **Warehousing Service**, through our tool “*Tetralogie*” and “*Xew*” [3]. **Big Data Analytics Service**, based on a specific Big Data Architecture [12, 13]. **Big Data Visualization Service**, the latter two services are located in a cloud, to achieve a real-time analysis.

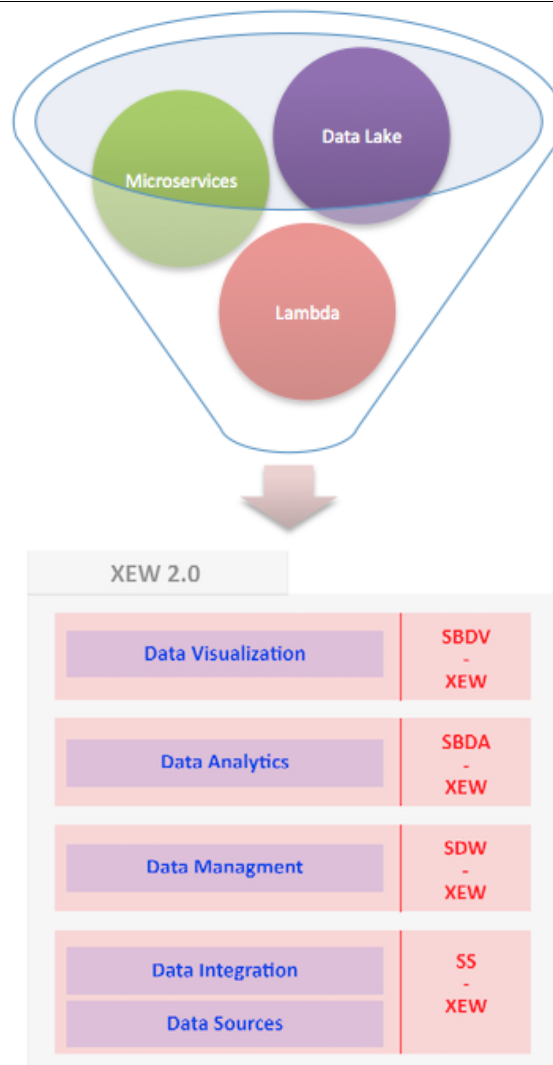


Figure 4. XEW 2.0 Architecture

3.2 Multilayer structure model of XEW 2.0

The concept of multilayer architecture has to address the problems of maintenance, reusability and scalability (or scaling) of an application. The Applications based on this architecture is divided into several logical levels, each with a clearly defined set of interfaces: the presentation layer, the data layer and one or more intermediate layers.

The middle layer (application logic) is simply constituted of code called by the user (via the presentation tier) to extract useful data. The presentation tier then receives the data and puts them into pages for their display. This separation of logic from and user interface significantly enhances the flexibility of designing an application. Indeed, it is now possible to create and run multiple user interfaces without changing the application logic, provided that the latter is endowed with a strictly defined communicate with the intermediate level.

This competitive intelligence tool uses intranet, internet, and various network programming languages to design and construct the information system platform. XEW 2.0 can be divided in four layers: top layer (user interface, web, 3G, 4G), middle layer (service, domain model), bottom layer (persistence), data layer (relational and multidimensional database) (cf. figure5).

The deployment of the platform it is based on the MVC pattern. To build an application using the following three levels:

- MODEL: is the core of the application. It maintains the state and the data that the application represents. When significant changes appear in the model, views are also changed.



- Controller: it defines the behavior of the application and manages the interaction with the user. It uses the model of components to meet user requests.
- VIEW: the user interface is the visual representation of the "model".

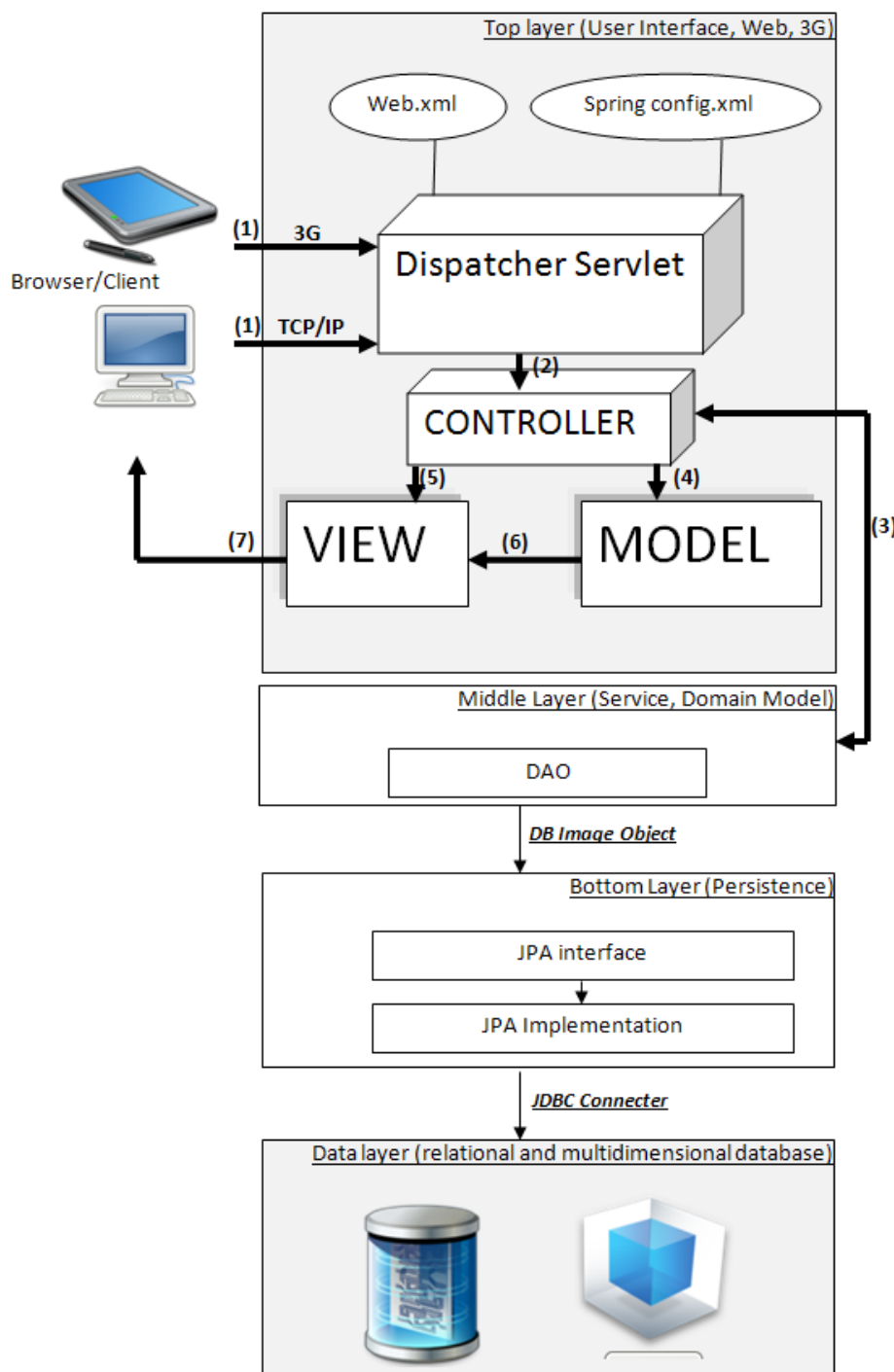


Figure 5. Multilayer structure of XEW 2.0

3.3 Monitoring, homogenization and structuring service: Warehousing service

In the proposed approach, we assume that the documents contained in the corpus can be target from heterogeneous sources. The heterogeneity of sources can consist in the format, language, etc... It is therefore to solve problems of semantic type (name clash or attribute type, no value, ...), but also structural type (untagged documents, HTML) or syntactic. To solve this problem, our approach is based on the principles of extracting information defined by [3]. This approach allows extracting predefined information from text documents, where



the location of information to be extract is marked or separated by strings. This solution allows us to: define an unified view of documents in the corpus target, and manage the various conflicts such as: “semantic, syntactic, inclusion and specificity”.

The unified view of the corpus corresponds with a structured, predefined, logical representation of collections sets. Its definition is based on the consideration of specific and generic format descriptors (cf figure 6) [4].

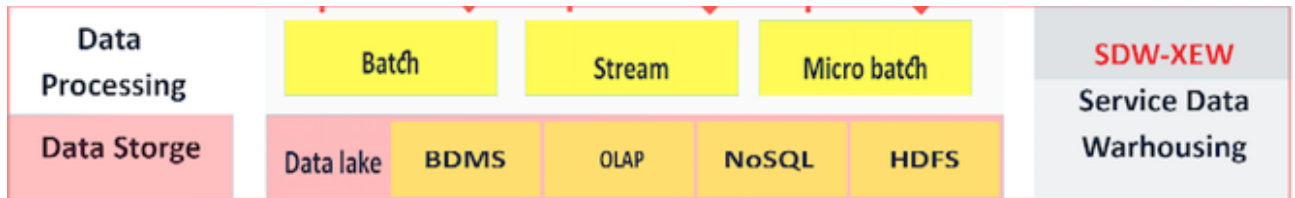


Figure 6. Information proceeding subsystem

3.4 Administration service and network structure

The need for security and management (administration network structure cf. figure 7) in a CI platform raises ham the strategic nature of information conveyed. Such security cannot be considered as an additional option that a CI platform can provides just in order to be distinguished from each another use. Especially as the leak of this information is not the result of inherent weaknesses in corporate computer systems, but above all it is an organizational issue.

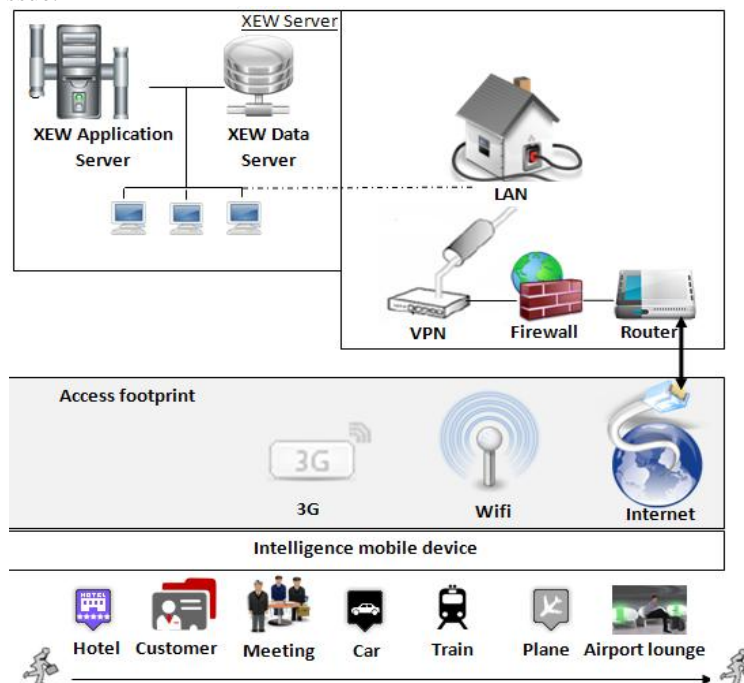


Figure 7. Network structure of XEW 2.0

So the issue is how to protect both data (CI as a product) and treatments (CI as a process). Thus, we propose a datum access control model and treatment in CI platforms [12, 13] where all tasks are only executed by authorized users and, thus enhance the privacy related with manipulation of data during the lifecycle of a CI process.

3.5 Reporting service

The « reporting » is the last and very important service to accomplish by CI process. In this level we propose 4 types of services: API XEW, Visualization Service, Real Time monitoring and CI decision. With these different services, it is possible to accede to strategic information anywhere. In order to ease the navigability on the strategic information, we intend to integrate specific visualization technique to each type of request like: evolutionary histograms, geographical charts, social networks, profile networks, semantic networks and the international networks (cf. figure 8).



Figure 8. Reporting service of XEW 2.0

4 Conclusion

In this paper, we present a new competitive intelligence system tools, dedicated to cover all stages of discovery, extraction and data management. With the evolution of technology, such a portal for CI in cloud will enables us to increase efficiency and responsiveness because at any moment, it is possible to gain access all strategic information by markers itself can be information back very quickly "field" which may possibly trigger other strategic analysis.

5 Bibliographie

- [1] Reverchon, A. (2015, 10 29). Big Data, la nouvelle frontière. Consulté le 06 03, 2017, sur PIXELS - Chroniques des révolutions numériques: www.lemonde.fr
- [2] Mongin, P., & Tognini, F. (2015). Petit manuel d'intelligence économique au quotidien 2ed: Comment collecter, analyser, diffuser et protéger son information. France: Dunod.
- [3] A. El Haddadi, Z. Boulouard, A. El Haddadi, A. Fennan, Conception Agile d'un Système d'Intelligence Economique, Séminaire Veille Stratégique Scientifique et Technique (VSST), Grenade, Espagne, Mai 2015.
- [4] A. El Haddadi, A. Fennan, A. El Haddadi, Z. Boulouard, L. Koutti, (2015), Mining unstructured data for a competitive intelligence system XEW. In: SIIIE 2015 – IEEE 6th International Conference on "Information Systems and Economic Intelligence, pp. 146–149.
- [5] SOSSON D., M. VASSARD, DOUSSET B., Portal for navigation in the strategic analysis , VSST'01, Vol 1, pp 347-358, Barcelone – Espagne, (2001).
- [6] GHALAMALLAH I., GRIMEH A., DOUSSET B., Processing data stream by relational analysis, EUROPEAN WORKSHOP ON DATA STREAM ANALYSIS, March, 14-16, CASERTA, ITALY, N° 36 , pp 67-70 (2007).
- [7] Guilhon, A., & Moinet, N. (2016). Intelligence économique : s'informer - se protéger - influencer. Paris: Pearson France.
- [8] Bao, Y.-l., & Hua, H.-F. (2015). Research on evaluation index system of tourist hotels' competitive intelligence ability. 4th International Conference on Computer Science and Network Technology (ICCSNT) (pp. 495 - 499). IEEE.
- [9] Gweth, G. (2015). 70 chroniques de guerre économique : 7 ans de veille et d'intelligence stratégique en Afrique. Collection Cnowdes.
- [10] Divya, S., Kriti, C., Preeti, V., & Kanika, J. (2015). Big Data - competitive intelligence. 2nd International Conference on Computing for Sustainable Global Development (INDIACom) (pp. 684 - 689). New Delhi, India: IEEE.
- [11] A. El Haddadi, A. El Haddadi, Z. Boulouard, F. Bouhafer, N. Chouati, A. Fennan, et al., (2017), Mining Unstructured Big Data for Competitive Intelligence and Business Intelligence. In: V. Sugumar, A.K. Sangaiah, A. Thangavelu (Eds.), Computational Intelligence and Applications in Business Intelligence and Big Data Analytics, 2017th ed., CRC Press, Taylor & Francis, Boca Raton, FL, pp. 327–348.
- [12] A. El Haddadi, A. El Haddadi, B. Dousset, A. Fennan, Le système d'intelligence économique - XEW. eTI, Revue électronique en Technologies de l'Information, Numéro 8, 2015.
- [13] Jing, D., Bo, H., & Wen, L. (2014). Research on Enterprise Competitive Intelligence Development and Strategies in the Big Data Era. IEEE International Conference on Computer and Information Technology (pp. 658 - 663). Xi'an, China: IEEE.