ITAIS 2021 Proceedings

Annual conference of the Italian Chapter of AIS (ITAIS)

2021

An Industrial Application of Business Intelligence Approach to the Electronic Defence Sector

Chiara Badolato
Elettronica S.p.A., Rome 00131, Italy, chiara.badolato@elt.it

Andrea Capuani
Elettronica S.p.A., Rome 00131, Italy, andrea.capuani@elt.it

Vittorio Concetti
Elettronica S.p.A., Rome 00131, Italy, vittorio.concetti@elt.it

Luigi Laura
Uninettuno University, Rome, Italy, luigi.laura@uninettunouniversity.net

Riccardo Rossi
Cy4Gate S.p.A. Rome 00131, Italy, riccardo.rossi@cy4gate.com

See next page for additional authors

An Industrial Application of Business Intelligence Approach to the Electronic Defence Sector

Chiara Badolato[1], Andrea Capuani[2], Vittorio Concetti[3], Luigi Laura[4], Riccardo Rossi[5] and Marco Smacchia[6]

[1]Elettronica S.p.A., Rome 00131, Italy
chiara.badolato@elt.it
[2]Elettronica S.p.A., Rome 00131, Italy
andrea.capuani@elt.it
[3]Elettronica S.p.A., Rome 00131, Italy
vittorio.concetti@elt.it
[4]Uninettuno University, Rome, Italy
luigi.laura@uninettunouniversity.net
[5]Cy4Gate S.p.A. Rome 00131, Italy
riccardo.rossi@cy4gate.com
[6]Luiss Guido Carli University, Rome 00197, Italy
msmacchia@luiss.it

Abstract. In the age of digital transformation, the availability of data is growing exponentially leading companies to struggle in processing big data while not missing out useful insights to focus on their business development strategy. In this scenario, always more often companies are making use of Business Intelligence platforms that could allow them to collect, analyses and disseminate data in real time to face the dynamic of the market. This paper aims to apply a Business Intelligence approach that adopts OSINT (open-source intelligence) and SOCMINT (Social Media Intelligence) techniques to Defence Electronics Market to analyse how this technology could facilitate Companies decision-making process by providing them with a distinct competitive advantage. In this frame we used QUIPO intelligence platform for an industrial scenario analysis in the Defence Electronics sector. This is an initial research to study the correlation between the experimental OSINT analysis carried out by the intelligence platform and the information based on the internal experience and know-how of the company for the use case study.

Keywords: Business Intelligence & Analytics, Defence, OSINT & SOCMINT, Digital Transformation.

1 Introduction

In Today's world, enterprises are under huge competitive pressures [1], markets are changing faster and faster and companies are subject to Digital Transformation challenges. According to McAfee [2], there is a need to adopt new data sources called Big Data which is defined as extremely large data sets in volume, velocity, variety, and veracity [3]. Big Data could help enterprises collect data and convert it into competitive advantages in the global market [2]. La Valle [4] states that the enterprises using data analytics to convert data into useful information outperform their competitors, by helping them understanding their business more deeply and improving the decision-making process [5]. As a consequence, data-driven enterprises such as Amazon, Microsoft, and Apple in less than ten years have replaced oil companies becoming the biggest firms for capitalization and profits worldwide[1]. Big Data is also getting bigger as information is coming from a larger number of devices and sources. In the last 2/3 years, humanity has generated more data than ever in its history[2]. It is no wonder that organization are forced to use only a selection of the vast amount of data available to be effective [6]. Considering such relevance of Big Data analysis, many companies must face constant challenges to keep up with digital transformation and become a data-driven company is one of them.

During this Digital Revolution, many companies rely on Business Intelligence (BI) software that can help analysts and management to more easily analyse and understand the primary and secondary data collected, facilitating the extraction of meaningful information for decision making.

Regarding BI Software, the two most important factors for its development and integration within a company are the creation of a meaningful data lake and correct taxonomic rules. According to Gibson [7, 8] without data there is nothing to analyse and collecting data in a wrong way could lead to inadmissible failures; so it is imperative to collect data for a right reason and in the right amount. The constitution of a data lake following precise data gathering rules is essential to increase the performance of the software, to considerably speed up the manual analysis of the collected or inserted documents, to limit the semantic engine errors when analysing unstructured data and, especially in OSINT research, to facilitate the process of assigning the items reliability. Due to the exponential increase in data, it is becoming increasingly difficult to monitor all the information around us. In this scenario, semantic analysis can help process and visualise online information and extract useful knowledge that helps people make informed decisions. For this to happen it is necessary to build new taxonomies each time the domain of interest changes [9]. To develop a new taxonomy, it is also essential to have a strong knowledge base on the specific domain [10]. The creation of structured taxonomic rules is fundamental to increase the software's analysis accuracy and speed up the examination of documents. Analysing the literature, the study of the impact of Business Intelligence is of paramount importance. More research is needed to verify the value of Business Intelligence, from both strategic and managerial aspects [3]. In this frame it will be fundamental to investigate the growing application of AI to BI platforms development [11].

In the first chapter, after introducing the main topics covered, we formalised the research question and conduct a literature review of the most important topics of the study: Business Intelligence and Open Source Intelligence. Chapters 2 and 3 deal

---

1 *Le prime 10 aziende al mondo negli ultimi 25 anni, Il sole 24 Ore 2018.* https://lab24.ilsole24ore.com/aziende-top/

2 A Day in Data, Raconteur. https://www.raconteur.net/infographics/a-day-in-data/

respectively with the research methodology and a study on the positioning of some Business Intelligence software within the market. In the fourth chapter is reported a case study on the application of the QUIPO Business Intelligence platform within the procurement department. After an introduction to the project, the structure of the platform and the crucial steps for its use are explained. Finally, the analysis is reported, and its results are discussed.

1.1 Purpose

This research is a starting point to study the validation process of a specific “use case” resulting from BI platform implementation, by answering, through the formulation of a conjecture, to the following research question: *"In the frame of the implementation of an OSINT Business Intelligence Platform Analysis in the defence market, does the creation of a meaningful Data Lake together with an appropriate Taxonomy structure make the experimental analysis results comparable and coherent to equivalent analysis validated by internal experience and know-how of the firm ?"*.

In this frame a segmentation of the BI Platform Market is performed in order to give an overview of the positioning of the QUIPO, a BI platform developed by Cy4Gate S.p.A.

Specifically, the aim is to verify the accuracy and reliability of business intelligence software based on artificial intelligence algorithms and machine learning, through the development of a case study that can provide empirical evidence on the subject. The experimental research conducted through the QUIPO platform will be compared with a validated research. The objective is to understand if the experimental research is able to give satisfactory results for the improvement of the decision-making process.

1.2 Business Intelligence Literature Review

There are different definitions of Business Intelligence; one of the oldest and most exhaustive is Hans Peter Luhn’s [12] a researcher at IBM: “an automatic system... developed to disseminate information to the various sections of any industrial, scientific, or government organisation”. Business intelligence is the process of transforming data into information, information into knowledge and knowledge into intelligence; with Artificial intelligence software, all this process is cyclical and automated. Karbhari [13] stated in his research that business intelligence is a broad category of application and technologies for collecting, storing, analysing and providing access to data to help the organisation to make a better decision. Chen et Al. [14] and Davenport [15] divide the evolution of BI software into three periods. The 1.0 phase begun in the 70s and concerns primarily the analysis of structured data with a focus on extraction, transformation and loading (ETL) processes to select decision-relevant data from transactional systems and bring them into the proper format for analyses [16]. To store these data, companies mostly used relational database management systems and data warehouses. Data were analysed mainly with the statistical method. With the spread of Internet in the 2000s, user-generated content and web analytics, collected through Web applications [17], are drivers for the second phase of BI. Business Intelligence 2.0 opened new frontiers for data analysis with more complex analysis techniques such as web and text mining and social network analysis [14]. With BI 3.0, we get into digital transformation, with data coming from mobile and IoT devices and the analysis of sensor-generated data [14]. With A Business intelligence software with OSINT (Open Source Intelligence), SOCMINT (Social Media Intelligence), AI tools, and limitless possibilities given by the Internet, we can access a large set of open information that provides us with a constant updating intelligence to improve decision making at all levels. Literature during these years has proven the benefits of BI, Torres et Al. [18] explain that Business intelligence helps companies in their Digital Transformation and improve organisational outcomes. BI also leads to efficiency improvement, process optimisation, time and cost reduction; taking in exam prior studies, we can also see an improvement in profitability, market share, and customer satisfaction [19]. According to Fink [19], BI assets generate value via two parallel mechanisms, operational and strategic, and these two dimensions not always are aligned but still lead to value creation. Business Intelligence also help the stakeholders identify the best solution for their firm and develop custom KPIs that best suit them [20].

1.3 OSINT Analysis Literature Review

In a 2011 document published by the Office of the Director of National Intelligence, OSINT was defined as: “intelligence produced from publicly available information that is collected, exploited, and disseminated in a timely manner to an appropriate audience for the purpose of addressing a specific intelligence requirement”.

OSINT is distinguished into two data types:

- Open-source data are published by individuals or groups without privacy restrictions. They have little value when taken individually but, when put together, can give a detailed overview of the situation being analysed.
- Open-source information (OSIF), on the other hand, includes all documents that can be obtained legally through request or purchase by citizens. They are usually more in-depth documents.

Through the use of OSINT tools, one can locate, extract and analyse documents from various sources such as blogs, social media, geolocation data, IP addresses, government documents etc. in order to produce intelligence [21]. OSINT analysis has had a lot of luck over the years because it is a quick and efficient way to carry out in-depth analysis [11]. It is widely used by many government agencies and in recent years also at corporate level. In fact, many companies that develop open-source intelligence tools are creating customised software for companies applications. OSINT is used to analyse links between various people, organisations, devices, and many other entities [22]. In the last two decades, there has been a shift from analysing to find hidden information to analysing to find relevant information [23]. The growth of data has created an environment where we are constantly flooded with information. It has become very difficult if not impossible for analysts to conduct manual searches without the help of this kind of software. In this scenario, artificial intelligence and machine learning algorithms are increasingly being used in Business Intelligence projects to partially manage various stages of the intelligence cycle. Although there are various interpretations of the intelligence cycle, which is commonly defined in 5 steps[3]:

- **Planning** the purpose and the activities of the analysis
- **Data Collection** to create the Data Lake
- **Processing** the documents gathered in the previous step (in AI-powered software, this stage is fully automated)
- **Analysis** of the documents
- **Dissemination** of the intelligence (Report or Graphic Dashboard)

These stages include the acquisition and validation of information, the identification of the value of the information and the distribution of the information to the client [7], [24, 25].

2 Methodology

In the first part of the research, we employed data from internal company databases and companies' websites to identify competitors to the QUIPO platform and to segment the market. In relation to the research objectives, we tried to segment the market through the construction of a two-variable positioning map. We examined a sample of 14[4] companies operating in the business intelligence market and we positioned them on the map according to their focus on the use of OSINT data and the possibility of customising taxonomies. The companies were chosen by referring to various reports concerning the business intelligence market. We selected market-leading BI platforms based on artificial intelligence and machine learning algorithms that simplify the decision-making process. An attempt was made to include a consistent number in order to represent as much of the market as possible. Consequently, we have analysed each platform individually to verify their correct positioning in the market by referring to two variables closely linked to the study. In the second part, we used an applied research method to empirically validate the use of the BI platform when applied to a specific "use case study". According to Kothari [26], Applied research *"aims at finding a solution for an immediate problem facing a society, or an industrial/business organisation, whereas fundamental research is mainly concerned with generalisations and with the formulation of a theory"*.

Specifically, we designed a use case within the company following the five phases of the intelligence cycle, paying particular attention to the data gathering phases for the data lake construction and the taxonomy development. Once the analysis was complete, we compared the data collected by the platform with the data from in-house intelligence and expertise to find a match between the experimental analysis performed by the QUIPO platform and the company's knowledge.

[3] The intelligence process/cycle (Source: "Joint Publication 2-0, Joint Intelligence". Defense Technical Information Center (DTIC). Department of Defense. February 2013)

[4] Note: Company names are not disclosed for confidentiality

3 Business Intelligence Market Segmentation

The two variables for the market segmentation have been selected by studying the literature on OSINT techniques and Business Intelligence. The dimensions relate to:

- **Focus on OSINT Analysis**, representing the degree to which the platform focuses on collecting and analysing external data available on the Internet.
- **Focus on TAXONOMY Customization**, represents the degree to which the platform focuses on taxonomies customisation according to the application cases.

Analysing the results deriving from the positioning map (see fig. 1)[5], we noticed the presence of four market niches corresponding to the quadrants outlined by the map plan.

Specific In-House. includes platforms that have a rigid semantic structure and that refer to internal non open-source data. These software collect data provided by the developing company or customer databases. They are used for specific purposes within the organisation that cannot be extended to other business functions. Since these are programs developed to solve a specific problem, the taxonomy structure is very rigid, focusing only on their scope of application such as CRM applications. The advantages of these systems come from their specialisation. As they are used for very specific tasks, they require less time to be implemented and benefit from more established taxonomies in their field of application.

Multi-Purpose In-House. includes platforms whose primary purpose is to use in-house data, but unlike the previous group, they have the ability to modify taxonomies depending on their field of application. They are used for various purposes and are very adaptable to various markets and various business functions, which is precisely the reason why they need to adapt the taxonomy structure depending on the type of analysis required. The advantages of these software lie in their greater elasticity of use, making them attractive in many sectors and for the most various problems.

Specific OSINT. includes platforms that focus on the collection and analysis of open-source data and have a limited ability to customise taxonomies for semantic analysis. Like the programs in the first quadrant, these software are also developed for very specific functions within organisations that require analysis of externally sourced data. The taxonomies are often fixed due to the high specificity of the task, even though platforms dedicated to OSINT research are very widely developed in their scope (e.g. this group includes cyber intelligence software). These platforms are able to manage the data they collect more quickly and accurately because of their knowledge of the primary sources from which they collect information and their highly developed semantic analysis around the main function.

[5] Note: Company names are not disclosed for confidentiality

Multi-Purpose OSINT. In this niche there are the Business Intelligence platforms that make open-source data collection their core business and allow developers and designers to create new taxonomies for semantic analysis depending on the platform's field of application. The tools in this quadrant have various features that make them suitable for various tasks, and their flexibility makes them usable by multiple functions within the company. Although they focus on OSINT research, these software are, in most cases, capable of handling internal data to combine the benefits of external and internal research. Since the strength of these platforms is their adaptability to various environments, they require ad hoc taxonomic development to reflect the client's needs. While implementation can be slightly time-consuming, once embedded in business operations, it provides a high degree of flexibility at all levels.

While aiming to help organisations in the collection, management, and analysis of data, BI platforms are very different from each other and can be used for different purposes. QUIPO, the BI Platform under study, is positioned within the Multi-Purpose OSINT industry market with its related key features to be considered for the adoption.

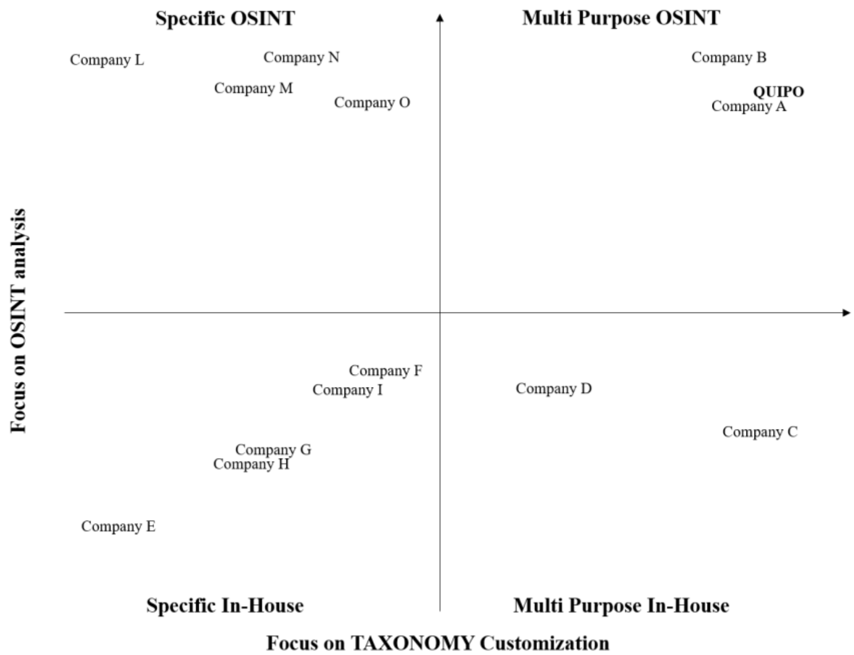

Fig. 1. Business Intelligence Positioning Map

4 QUIPO Application Case

The objective is to evaluate the benefits and opportunities that OSINT analysis can bring to the Company organisation by creating several use cases. This study is part of the company's project to integrate QUIPO into the procurement function in order to achieve the following objectives:

- Improvement of the effectiveness of the suppliers scouting process
- Simplification of the suppliers "due diligence" (actual or potential).
- Technological development trend of specific product categories, by identifying leading edge or out-of-market suppliers.

During our research, we followed the intelligence cycle described above by focusing on product capability to determine if the information we found during the analysis of the documents collected by the platform were correlated to the information held in the company. In the use case, we have examined three potential suppliers, scouting their capabilities to explore potential collaboration opportunities.

4.1 QUIPO Structure

QUIPO consists of three main modules linked together by tools that exploit artificial intelligence and machine learning algorithms to analyse documents:

- The first module consists of a non-relational database with the task of indexing documents from internal and external sources. An archive is then created containing all the documents, enriched with the first metadata that establishes their uniqueness and allows the internal search engine to retrieve them quickly. After being saved and indexed in the database, the data is analysed by the platform's semantic engine or, in the case of images, by image and face recognition tools. In this phase, the documents are assigned additional metadata, which is very important for the representation and analysis by the user.

- The second module deals with the representation of the documents collected in the database using the metadata produced in the previous phases. Through the aggregation of this data, the module allows the creation of dashboards and graphs with which we have obtained the data to formulate our conjecture.

- The third module is the platform's GUI, which allows the end-user to manage all the tasks necessary to carry out the analysis. It can be considered as the point of contact between the backend tools and the analyst. Closely linked to this module is the data dissemination platform that allows the textual and graphical display of the work done.

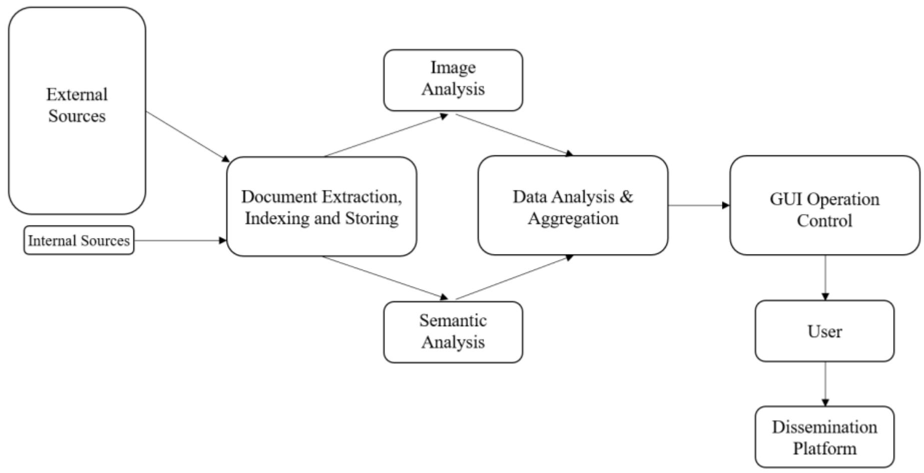

Fig. 2. Visual representation of Quipo Structure

4.2 Data Lake Construction

The data-gathering phase is one of the most delicate moments of this kind of analysis. In this step it is essential to perform a first skimming of the collected articles, doing searches as relevant and intelligent as possible. In fact, the risk is to conduct summary searches for keywords with very broad meanings and find the archive filled with tens of thousands of documents that are completely useless for the analysis by increasing the analyst's difficulty in recognising the articles that must be included in the task. At this stage, it is necessary to have a thorough knowledge of the platform and its reasoning method for data ingestion. In this way, it is very easy to make more detailed research that allows having a clearer overview of the phenomenon observed at first glance.

In the construction of the data lake, we source news from major defence websites and magazines, Google News, Europe Media Monitor, the press releases of the analysed companies and related organisations, ministerial websites, social media (Twitter and Facebook), the press release, archives of documents already present in the platform from previous searches.

In choosing the sources and methods of data gathering applied, we respected the NATO Open-Source Intelligence Handbook guidelines [7, 8]:

1. The authority of the source
2. The accuracy (by validating it against other sources)
3. The objectivity of the source
4. The currency (i.e., the provision of a timestamp for publication and the presence of an author)
5. The coverage (the degree of relevancy)

The data gathering process was conducted by searching for very precise keywords that included the name of the companies under study, their products or geopolitical concepts related to the research. In three months, we had 30,000 articles in our specific archive, to which we added other archives containing news from magazines and defence fairs on the major players in the market (500,000), the press release (200,000), news on investment banks and related companies (20,000) and geopolitical information (1,500,000). Our data lake, at the time of analysis, consisted of 2,250,000 articles analysed by the platform. The next step was to process them by filtering the documents within the data lake according to taxonomic concepts and the companies we were interested in. In the selection of the significant articles, the same rules for the identification of sources were applied. It was necessary to repeat the process because some media such as Google News, Facebook and Twitter require more careful analysis to avoid the ingestion of unreliable news. After filtering the articles of interest, reliability (low, medium, high) was assigned to the most interesting ones. Finally, an archive containing all these documents was created and compared with the attainable information in our possession.

4.3 Taxonomy Creation

A taxonomy is a semantic hierarchy that organises concepts by is-a relations, which exhibits the capability of improving many NLP task [10]. Its construction is one of the most important operations in the implementation of software capable of semantic analysis. A taxonomy is composed of rules that allow the semantic engine to understand a given taxonomic entity within a document. A taxonomy development depends on the domain that has to be analysed and the user requirement, who may have different needs according to his objective. A taxonomy can be created using two types of approach:

- Categorisation: a deductive method in which the starting point is a general concept to arrive at specific individual entities. In this approach, a semantic tree is constructed with branches comprising subcategories of more general entities. Categories are generated first by formulating their conceptual descriptions and then classifying objects according to the descriptions.
- Extraction: The inductive method consists of identifying, aggregating, and normalising specific semantic entities to arrive at more generic concepts.

After a knowledge analysis on the reference domain and the choice of the approach to be used, rules are developed and associated to a single domain. These rules allow the platform to be able to distinguish ambiguous words and concepts within a context, making data analysis more precise.

In our analysis, we needed to initially develop an entry-level taxonomy of Electronic Warfare (EW). Our objective was to contextualise the possible suppliers within the market, identifying their capabilities and heritage. Therefore, a categorisation approach was used to create a semantic tree composed of general concepts that helped us understand the scope in which a company operates and specific entities to study the differences between players. The taxonomy was validated using a benchmark test to establish the correctness of taxonomic rules. Through the use of already analysed libraries, the test can give a percentage on the accuracy of the taxonomy. It is based on an iterative process that allows the rules to be evaluated and modified according to the result.

5 Proceedings and Results

For this research, we focused on the first cluster of taxonomies available in the procurement project concerning firm Electronic Warfare capabilities in the defence sector, by examining three firms that belongs to the same capability cluster. We inspected several reports available in the company and compiled a file containing information about the contracts, merge & acquisition activities, partnership and joint ventures of the organisations under study. The file containing 170 occurrences was then validated by the company functions. At the same time, we analysed the documents contained in the QUIPO data lake, resulting in a final archive of 932 items. We selected the nine taxonomic concepts concerning Electronic Warfare capabilities within the documents and divided into clusters all the articles for each specific concept. We then calculated the percentage of each cluster with respect to the total number of documents in the archive to normalise the data. Next, we studied the existing correlation between the two matrices through the following formula for the calculation of the Pearson Correlation Coefficient:

$$r = \frac{\sum[(x_i - \bar{x})(y_i - \bar{y})]}{\sqrt{\sum(x_i - \bar{x})^2 * \sum(y_i - \bar{y})^2}}$$

The results (see fig.3, table 1) indicate a highly positive correlation [27] between the two matrices ($r = 0.927$, $R2 = 0.86$). That denotes a correspondence between analysis based solely on OSINT techniques by using a business intelligence platform and data validated by firm's experts. An empirical correlation also emerged when compiling the report, noting a sufficient match between the concepts extracted from the platform and those present in the reports analysed.

In light of the following results, we can formulate the following conjecture:

"In the development of a Business Intelligence platform, which uses analysis techniques based on the OSINT intelligence cycle, the creation of a data lake following precise rules that point to correct use of sources together with the creation of an appropriate taxonomy to the specific domain, allows the number of occurrences on a given entity to be very close to reality".

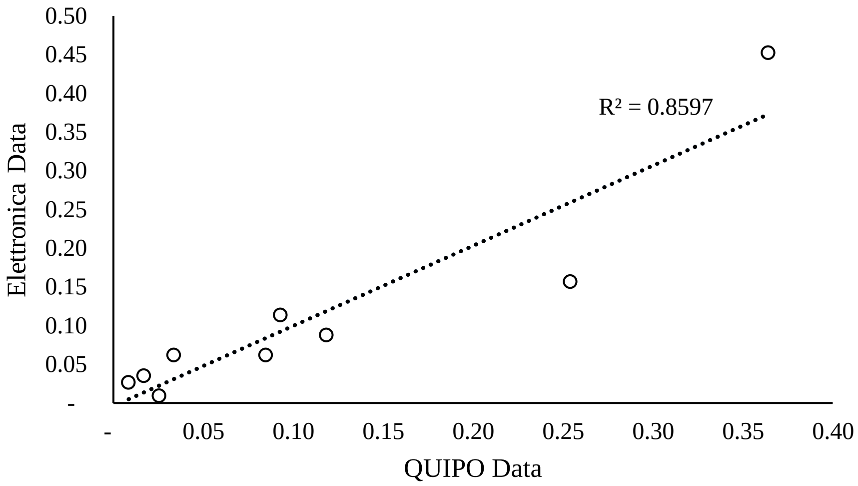

Fig. 3. Graphical representation of the correlation between the two matrices.

Table 1. Capabilities examined and their percentages on total documents.

Capabilities	QUIPO Data	ELT Data
C4ISR	0,36	0,45
DIRCM	0,09	0,11
Training	0,25	0,16
Customer Support	0,03	0,06
RECM	0,01	0,03
COMINT	0,12	0,09
RESM	0,02	0,03
SIGINT	0,08	0,06
ELINT	0,03	0,01

6 Conclusion, limitations and Future Work

As a result of the digital transformation development, the concept of Business Intelligence is increasingly associated with artificial intelligence and machine learning. More and more organisations worldwide are adopting these systems, profoundly changing their mindset. It is therefore of paramount importance to empirically study the results of applying these technologies in various fields and for various purposes. We are interested in giving a practical demonstration of the application of a new BI technology to support decision-making within an organisation. Furthermore, this study contributes to the literature by describing techniques for implementing a business intelligence application.

In the following research, we have described a technique for the construction of a data lake and the creation of a taxonomy in a specific domain through the use of a practical case that may be useful for future case studies. We have also seen how the information extracted by the BI platform, based on OSINT analysis techniques, in the frame of the use case analysis is correlated to the information available and validated by company expertise. The data extracted in the experimental analysis conducted through open-source channels was able to identify electronic warfare information very similar to company research. Accuracy is very important as this software could save a lot of time in decision-making. In fact, once implemented, they continue to analyse data over time giving continuous updates on the analysed topic. The study was carried out with reference to a single taxonomy and within a specific domain, future studies could extend the interest to different application domains to deepen their dynamics of operation and the results derived from their implementation. In addition, could be very important to develop new metrics to measure certain results in the medium to long term. This study is only the starting point for a broader project, aiming to expand the platform's scope through the creation of various taxonomies in the procurement domain and to continue research into the application of Business Intelligence platforms powered by AI tools in areas not yet studied.

References

1. Knight, Gary A., and Peter W. Liesch: Internationalisation: From incremental to born global. Journal of World Business 51 (1): 93–102 (2016).
2. McAfee, A., E. Brynjolfsson, T.H. Davenport, D.J. Patil, and D. Barton: Big data: The management revolution. Harvard Business Review 90 (10): 60–68 (2012).
3. Liang, T.P., and Y.H. Liu: Research landscape of business intelligence and big data analytics: A bibliometrics study. Expert Systems with Applications 111: 2–10 (2018).
4. LaValle, S., E. Lesser, R. Shockley, M.S. Hopkins, and N. Kruschwitz: Big data, analytics and the path from insights to value. MIT sloan management review 52 (2): 21 (2011).
5. Ahi, A., G. Baronchelli, O. Kuivalainen, and M. Piantoni. International market entry: How do small and medium-sized enterprises make decisions? Journal of International Marketing 25 (1): 1–21 (2017).
6. Thomas C. Redman, Does your Company Know What to Do with All Its Data? Harvard Business review. https://hbr.org/2017/06/does-your-company-know-what-to-do-with-all-its-data (2017).
7. Babak, Akhgar, P. Saskia, Bayerl, Fraser, Sampson, Open-Source Intelligence Investigation From Strategy to Implementation. 1st edn. Springer International Publishing AG (2016).
8. Gibson H., Ramwell S., Day T. (2016) Analysis, Interpretation and Validation of Open Source Data. In: Akhgar B., Bayerl P., Sampson F. (eds) Open Source Intelligence Investigation. Advanced Sciences and Technologies for Security Applications. Springer, Cham. https://doi.org/10.1007/978-3-319-47671-1_7
9. Carrion, B., Onorati, T., Díaz, P. et al. A taxonomy generation tool for semantic visual analysis of large corpus of documents. Multimed Tools Appl 78, 32919–32937 https://doi.org/10.1007/s11042-019-07880-y (2019).

10. Chengyu Wang, Xiaofeng He, Aoying Zhou: A Short Survey on Taxonomy Learning from Text Corpora: Issues, Resources and Recent Advances. Conference on Empiri-cal Methods in Natural Language Processing, pages 1190–1203 (2017).
11. João Rafael Gonçalves Evangelista, Renato José Sassi, Márcio Romero & Domingos Napolitano: Systematic Literature Review to Investigate the Application of Open Source Intelligence (OSINT) with Artificial Intelligence, Journal of Applied Security Research, DOI: 10.1080/19361610.2020.1761737 (2020).
12. Luhn H.P., A business intelligence system IBM Journal of Research and https://doi.org/10.1147/rd.24.0314 (1958).
13. P. Karbhari, "SOA trends: Service oriented business intelligence," White Paper. [Online]. Available: www.technologyexecutivesclub.com (2006).
14. Chen, H., R.H.L. Chiang, and C.S. Veda: Business intelligence and analytics: from big data to big impact. MIS Quarterly 36 (4): 1165–1188. https://doi.org/10.2307/41703503 (2012).
15. Davenport, T.H. Analytics 3.0: In the new era, big data will power consumer products and services. Harward Business Review 91: 64–72 (2013).
16. Chaudhuri, S., U. Dayal, and V. Narasayya: An overview of business intelligence technology. Communications of the ACM 54 (8):88. https://doi.org/10.1145/1978542.1978562 (2011).
17. O'Reilly, T.: What Is Web 2.0? Design Patterns and Business Models for the Next Generation of Software. https://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html (2020).
18. Torres,R. Sidorova,,A. Jones,M: Enabling firm performance through business intelligence and analytics: A dynamic capabilities perspective. Information & Management 1 (2019) 1–18 (2018).
19. Fink,L. Yogev,N. Even: A Business intelligence and organisational learning: An empirical investigation of value creation processes. Information & Management 54 38–56 (2017).
20. Tripathi, Anuj and Bagga, Teena, Leading Business Intelligence (BI) Solutions and Market Trends: Proceedings of the International Conference on Innovative Computing & Communications (ICICC) 2020, Available at SSRN: https://ssrn.com/abstract=3568414 or http://dx.doi.org/10.2139/ssrn.3568414 (April 4, 2020).
21. Quick, D., & Choo, K. K. R.: Digital forensic intelligence: Data subsets and Open Source Intelligence (DFINTþOSINT): A timely and cohesive mix. Future Generation Computer Systems, 78, 558–567. https://doi.org/10.1016/j.future.2016.12.032 (2018).
22. Maciołek, P., & Dobrowolski, G.: Cluo: Web-scale text mining system for open source intelligence purposes. Computer Science, 14(1), 45–62. https://doi.org/10.7494/csci.2013.14.1.45 (2013).
23. Nicart, E., Zanuttini, B., Gilbert, H., Grilheres, B., & Praca, F.: Building document treatment chains using reinforcement learning and intuitive feedback [Paper presentation]. 2016 IEEE 28th International Conference on Tools with Artificial Intelligence (ICTAI) (pp. 635–639). IEEE. https://doi.org/10.1109/ICTAI.2016.0102 (2016).
24. Heather J. Williams, Ilana Blum: Defining Second Generation Open Source Intelligence (OSINT) for the Defense Enterprise. RAND National defense Research Institute ISBN: 978-0-8330-9883-2 (2018).
25. Marzell, Laurence: OSINT as Part of the Strategic National Security Landscape. 10.1007/978-3-319-47671-1_4 (2016).
26. Kothari, C. R., Research Methodology, Methods and Techniques (2nd ed., pp. 109-110). New Delhi: New Age Inter- national (P) Limited (2008).
27. M.M Mukaka, Statistics Corner: A guide to appropriate use of Correlation coefficient in medical research Malawi Medical Journal; 24(3): 69-71 September (2012).