ISSN: 2001-015X

Journal of Intelligence Studies in Business

Vol. 8, No. 2 2018

Included in this printed copy:

Editor-in-chief:
Klaus Solberg Søilen

Journal of Intelligence Studies in Business

Vol 8, No 2, 2018

The **Journal of Intelligence Studies in Business** (JISIB) is a double-blind peer reviewed, open access journal published by Halmstad University, Sweden. Its mission is to help facilitate and publish original research, conference proceedings and book reviews.

FOCUS AND SCOPE

The journal includes articles within areas such as Competitive Intelligence, Business Intelligence, Market Intelligence, Scientific and Technical Intelligence and Geo-economics. This means that the journal has a managerial as well as an applied technical side (Information Systems), as these are now well integrated in real life Business Intelligence solutions. By focusing on business applications, this journal does not compete directly with the journals that deal with library sciences or state and military intelligence studies. Topics within the selected study areas should show clear practical implications.

OPEN ACCESS

This journal provides immediate open access to its content on the principle that making research freely available to the public supports a greater global exchange of knowledge. There are no costs to authors for publication in the journal. This extends to processing charges (APCs) and submission charges.

COPYRIGHT NOTICE

PUBLICATION ETHICS

The journal's ethic statement is based on COPE's Best Practice Guidelines for Journal Editors. It outlines the code of conduct for all authors, reviewers and editors involved in the production and publication of material in the journal. An unabridged version of the journal's ethics statement is available at https://ojs.hh.se/.

Publication decisions: The editor is responsible for deciding which of the articles submitted to the journal should be published. The editor may be guided by the policies of the journal's editorial board and constrained by such legal requirements as shall then be in force regarding libel, copyright infringement and plagiarism. The editor may confer with other editors or reviewers in making this decision. *Fair play*: An editor will evaluate manuscripts for their intellectual content without regard to race, gender, sexual orientation, religious belief, ethnic origin, citizenship, or political philosophy of the authors. *Confidentiality:* The editor and any editorial staff must not disclose any information about a submitted manuscript to anyone other than the corresponding author, reviewers, potential reviewers, other editorial advisers, and the publisher, as appropriate. *Disclosure and conflicts of interest*: Unpublished materials disclosed in a submitted manuscript must not be used in an editor's own research without the express written consent of the author.

Duties of Reviewers
Promptness: Any selected referee who feels unqualified to review the research reported in a manuscript, is aware of a personal conflict of interest, or knows that its prompt review will be impossible should notify the editor and excuse himself from the review process. *Confidentiality*: Any manuscripts received for review must be treated as confidential documents. *Standards of Objectivity*: Reviews should be conducted objectively. Referees should express their views clearly with supporting arguments. *Acknowledgement of Sources*: Reviewers should identify relevant published work that has not been cited by the authors. *Disclosure and Conflict of Interest*: Privileged information or ideas obtained through peer review must be kept confidential and not used for personal advantage.

Duties of Authors
Reporting standards: Authors of reports of original research should present an accurate account of the work performed as well as an objective discussion of its significance. Fraudulent or knowingly inaccurate statements constitute unethical behavior and are unacceptable. *Data Access and Retention*: Authors are asked to provide the raw data in connection with a paper for editorial review, and should be prepared to provide public access to such data (consistent with the ALPSP-STM Statement on Data and Databases). *Originality and Plagiarism*: The authors should ensure that they have written entirely original works, and if the authors have used the work and/or words of others that this has been appropriately cited or quoted. *Multiple, Redundant or Concurrent Publication*: An author should not publish manuscripts describing essentially the same research in more than one journal or primary publication. Submitting the same manuscript to more than one journal concurrently constitutes unethical publishing behaviour and is unacceptable. *Acknowledgement of Sources*: Proper acknowledgment of the work of others must always be given. *Authorship of the Paper*: Authorship should be limited to those who have made a significant contribution to the conception, design, execution, or interpretation of the reported study. The corresponding author should ensure that all appropriate co-authors and no inappropriate co-authors are included on the paper, and that all co-authors have seen and approved the final version of the paper and have agreed to its submission for publication. *Disclosure and Conflicts of Interest*: All authors should disclose in their manuscript any financial or other substantive conflict of interest that might be construed to influence the results or interpretation of their manuscript. All sources of financial support for the project should be disclosed. *Fundamental errors in published works*: When an author discovers a significant error or inaccuracy in his/her own published work, it is the author's obligation to promptly notify the journal editor or publisher and cooperate with the editor to retract or correct the paper.

ARCHIVING

This journal utilizes the LOCKSS system to create a distributed archiving system among participating libraries and permits those libraries to create permanent archives of the journal for purposes of preservation and restoration.

PUBLISHER

Halmstad University, Sweden
First published in 2011. ISSN: 2001-015X.
Owned by Adhou Communications AB

Journal of Intelligence Studies in Business

Stevan Dedijer
1911-2004

JISIB
Journal of Intelligence Studies in Business
Vol. 8, No 2 (2018) p. 4-5
Open Access: Freely available at: https://ojs.hh.se/


Stevan Dedijer
1911-2004

EDITOR'S NOTE **VOL 8, NO 2 (2018)**

Social media intelligence

Web-intelligence is not new to intelligence studies. IT plus intelligence has been the most frequent topic of the issues published the past years. Thus Vol 8, No 1 (2018) is entitled "The disciplines of management and IT have indeed merged: new empirical data", Vol 7, No 1 (2017) "Business intelligence, big data and theory" and Vol 6, No 3 (2016) "What role does technology play for intelligence studies at the start of the 21st century?". Special issues have looked at the problem of IT failures in relation to business intelligence: "How companies succeed and fail to succeed with the implementation of intelligence systems", Vol 7, No 3 (2017) and "How companies work and fail to work with business intelligence, Vol 7, No 2 (2017). During the past years companies have indeed learned from their failures. Maybe this phase was inevitable as a part of growing up. We see the same development on e-commerce sites: they mostly work well now, but didn't just a few years ago. A certain difference between countries still exists, but the industry is getting there. Closely related to failures of implementation are user perspectives on business intelligence systems, which have resulted in numerous research articles. A well-cited article by Adamala and Cidrin (2011) led to the development of several models and theories as presented, for example, in Vol 6, No 2 (2016) entitled "User perspectives on business intelligence".

The focus in JISIB is always technology. It is more a question of which aspect of technology we focus on. In this issue, it is social media or social media intelligence. The paper by Gioti and Ponis entitled "Social business intelligence: Review and research directions" is a literature review exploring the new direction of social business intelligence (SBI), where social media meets BI. The last paper is entitled "Business intelligence for social media interaction in the travel industry in Indonesia". The authors, Yulianto, Girsang and Rumagit propose a way to develop a data warehouse to analyze data from social media, such as likes, comments and sentiment, applied to the travel industry in Indonesia.

Another aspect of the journal maintains the tradition of intelligence studies in general. Intelligence studies must always be broad to be relevant and not to miss important pieces. Specialization is a necessity and a curse at the same time. Vol 6, No 1 (2016) in entitled "The width and scope of intelligence studies in business". A part of this width and critique has involved self-reflection. Thus earlier articles in JISIB often discussed methods. Case studies (by country or industry) were always a favorite. In Vol 4, No 3 (2014) JISIB continued this tradition of publishing case studies. In Vol 3, No 2 (2013), the whole issue is dedicated to one country; Brazil. Analyzing patents analysis has also been a frequent and reoccurring topic. In this issue both of these directions are represented. The third article is entitled "Investigating the competitive ıntelligence practices of Peruvian fresh grapes exporters," written by Bisson, Mercedes, and Tong. The authors suggest a number of changes for Peruvian grapes exporters to become more competitive based on a CI approach.

The fourth paper by Shaikh and Singhal entitled "An analysis of ip management strategies of ict companies based on patent filings" tries to identify the strategies of five US and Indian IT companies by analyzing their patents. The first paper by Nuortimo is entitled "Measuring public acceptance with opinion mining: The case of the energy industry with long-term coal R&D investment projects" and is part of his dissertation in science communication at the Faculty of Humanities at the University of Oulu. The paper shows how opinion mining can be used effectively, and was one of a series presented at the ICI Conference in Bad Nauheim this year. Many of the earlier papers in JISIB came directly from academic or practitioners' conferences. In Vol 2, No 1 (2012) it said: "The journal works in symbioses with a number of conferences. It relies heavily on the contributions of scientific papers presented at these conferences, in particular for these first issues. Among these we would in particular like to mention the more scholarly conferences, like VSST, ECIS, ICTICTI and SIIE. In the near future we also hope to receive contributions from INOSA and ECKM. We also receive support from members in the more professional conferences related to Intelligence Studies like ICI and SCIP" (p. 4). And Vol 3, No 3 (2013): "The journal continues to draw mainly on articles presented at academic conferences on topics related to competitive intelligence. In 2013 SCIP organized a first conference in South Africa, under the leadership of ASA du Toit, the journal's editor for Africa.". And in Vol 2, No 3 (2012): "Most contributions continue to come from the best papers from a number of conferences related to Intelligence Studies. Two out of five articles come from ECKM 2012, which was held 6-7 September in Cartagena, Spain." And in Vol 2, No 2 (2012) echoed a similar sentiment. Today the number of conferences has been reduced for different reasons, which it takes too long to get into here and now.

The last group of articles worth mentioning is opinion pieces. These are non-empirical articles. Today they are less frequent, but at the beginning they served another role, as pointed out in Vol 4, No 1 (2014): "In this issue of JISIB we have admitted a large number of opinion pieces. Opinion pieces are important to allow for a broader perspective of the field in terms of policies, adaptions of CI in foreign countries and general interest in the form of debates. It also shows the normative qualities that are present in any social science discipline". At the very beginning it was also made clear that the goal was always to be relevant for practitioners. Thus in Vol 1, No 1 (2011) we read: "The final aim of the journal is to be of use to practitioners. We are not interested in theory for the sake of theory, and we do not want to publish solutions to small problems which will have no real impact in the intelligence field.". With your help we try to keep with that goal.

As always, we would above all like to thank the authors for their contributions to this issue of JISIB. Thanks to Dr. Allison Perrigo for reviewing English grammar and helping with layout design for all articles and to the Swedish Research Council for continuous financial support. A special congratulation goes to Rainer Michaeli for having taken the ICI conference to its 10th anniversary. Well done, and thank you for the ongoing cooperation.

On behalf of the Editorial Board,

Sincerely Yours,

Prof. Dr. Klaus Solberg Søilen
Halmstad University, Sweden
Editor-in-chief

JISIB

Journal of Intelligence Studies in Business
Vol. 8, No. 2 (2018) pp. 6-22
Open Access: Freely available at: https://ojs.hh.se/

Stevan Dedijer
1911-2004

Measuring public acceptance with opinion mining: The case of the energy industry with long-term coal R&D investment projects

Kalle Nuortimo[a*]

[a] *Sumitomo SHI FW Energia Oy, P.O.Box 201, FIN-78201, Varkaus, Finland*

Corresponding author (*): *kalle.nuortimo@shi-g.com*

Received 18 June 2018 Accepted 21 August 2018

ABSTRACT New Web 2.0-based technologies have emerged in the field of competitor/market intelligence. This paper discusses the factors influencing long-term product development, namely coal combustion long-term R&D/Carbon Capture and Storage (CCS) technology, and presents a new method application for studying it via opinion mining. The technology market deployment has been challenged by public acceptance. The media images/opinions of coal power and CCS are studied through the opinion mining approach with a global machine learning based media analysis using M-Adaptive software. This is a big data-based learning machine media sentiment analysis focusing on both editorial and social media, including both structured data from payable sources and unstructured data from social media. If the public acceptance is ignored, it can at its worst cause delayed or abandoned market deployment of long-term energy production technologies, accompanied by techno-economic issues. The results are threefold: firstly, it is suggested that this type of methodology can be applied to this type of research problem. Secondly, from the case study, it is apparent that CCS is unknown also based on this type of approach. Finally, poor media exposure may have influenced technology market deployment in the case of CCS.

This paper is the extended version of a paper from the ICI 2018 international conference on Competitive & Market intelligence, June 5-8 Bad Neuheim, Germany.

KEYWORDS Carbon Capture and Storage, CCS, greenhouse gas control, market deployment, opinion mining, public acceptance, web-intelligence

1. INTRODUCTION: EMERGING WEB-INTELLIGENCE APPLICATIONS FOR COMPETITOR AND MARKET INTELLIGENCE

The aim of competitive intelligence (CI) is to analyse and exploit information about a company's competitors and sectors of activity to determine its competitive strategy and to develop new knowledge about its competitors in an increasingly complex and fast-moving economy to maintain levels of innovation and thus gain a competitive advantage (Amarouche et al. 2015). The most popular term used in the literature is competitive intelligence, followed by business intelligence (BI) and market intelligence (MI) (Dutoit 2015).

The lack of sufficient and reliable information sources about competitors can restrict the capability of CI (Xu et al. 2010). Traditionally, information about competitors has mainly been obtained from press releases, analyst reports, and trade journals, and recently also from competitors' websites and news sites. Unfortunately, such information is mostly generated by the company that produces the product; therefore the amount of information is limited and its objectivity is questionable (Xu, et al. 2010). Competitive intelligence is favoured at the expense of strategic management as a field and has evolved over the years as a result of the need for enterprises to scan the complex external environment (Dutoit 2015). Competitive intelligence provides the company with a clearer picture of its competitive environment, while the increasingly frequent use of information and communication technologies (ICT), including online shopping sites, blogs, social network sites, and forums, provides incentives for companies to promote their advantages over their competitors (Amarouche et al. 2015).

Due to the emergence of Web 2.0, including social media, CI now has a potentially wide field for developing new applications. The large numbers of customer-generated product reviews often contain information about competitors and have become an interesting source of competitive and market intelligence to mine (Xu, et al. 2010). Finding the weakness of products from customer feedback can help manufacturers improve their product quality and competitive strength. In recent years, more and more people have begun expressing their opinions about products online, and both the feedback of manufacturers' own products and their competitors' products could be easily collected (Chang et al. 2012).

Several applications have been developed for next generation CI/MI. The opportunities associated with data and analysis in different organizations have helped generate significant interest in business intelligence and analysis (BI&A). BI&A is often described as the techniques, technologies, systems, practices, methodologies, and applications for analysing critical business data to help an enterprise better understand its business and market, and to make timely business decisions (Chen et al. 2012). Opinion mining in product CI was discussed by Amarouche et al. (2015). A system to efficiently analyse patent data, a patent trend change mining (PTCM) approach that can identify changes in patent trends without the need for specialist knowledge, has been proposed by Shih et al. (2010). Market intelligence from microblogs, which have become great sources of consumer opinions, has been developed in the form of compact numeric summarization of opinions by Li et al. (2013), from which the proposed mechanism can effectively discover market intelligence (MI) to support decision-makers. In 2012, Chang et al. introduced Weakness Finder, which helps manufacturers find their product weakness by using aspect-based sentiment analysis on Chinese reviews. In computational linguistics, irony is one of the more challenging topics in sentiment classification, and tools to detect irony were described by Reyes and Rosso in research focusing on identifying key components for the task of irony detection (2012).

This paper describes an opinion mining approach to discover the public acceptance of carbon capture and storage (CCS) technology, in order to highlight influences on long-term R&D strategy. Compared to media images of solar and biomass power (Nuortimo 2017a&b), differences exist, and can be used to highlight the link and differences between existing theoretical base.

2. CASE CARBON CAPTURE AND STORAGE (CSS)

The need to reduce atmospheric CO_2 has resulted in several global agreements (e.g. Kyoto Protocol, 1997; Paris Agreement, 2015), all affecting environmental legislation, technology strategies, and decision-making of individual companies. The large-scale adoption of CCS in combination with increased energy efficiency is seen as one option to halt CO_2 emissions in the short run (Wennersten et al. 2015). Power plants with CCS in addition to large shares of low carbon generators such as renewables would be required to meet the global targets (Brouwer et al. 2015). Carbon capture and storage facilities coupled with energy efficient power plants would provide a strategy to permit the continued use of fossil fuels whilst reducing CO_2 emissions. The CCS process includes three stages of capture and compression of CO_2 from power stations, transport of CO_2, and storage away from the atmosphere for hundreds to thousands of years (Hammond et al. 2011).

However, regardless of the potential, the technology deployment has not been realised due to lack of economic incentives, regulations, and public acceptance (Nuortimo 2012). Technologies have been connected with societal controversies in the past; for example, nuclear power and gene technologies have been surrounded by dispute, potentially causing public rejection. Past rejection of technologies by the public emphasises the urgency to understand the psychological features of societal acceptance of technologies (Gupta *et al.* 2012). Public acceptance of technologies such as CCS is crucial for successful introduction into the society (Huijts et al. 2012). In this study, the media image of CCS, especially in social media (SoMe), was studied to find possible implications for public acceptance of CCS technology. This was done by reviewing the relevant CCS discussions and studying the media image of CCS from 2014 to 2016. The main research question is formulated as: what is the media image of CCS and its possible implications for public acceptance, and, furthermore, how does this relate to coal combustion technologies in general?

This paper is organised as follows. First the literature is analysed in terms of the important aspects of CI/MI tools and developments, case CCS and related public acceptance and market deployment, and subsequently with application of the new method, opinion mining with machine-based media analysis. A possible link from media image to product market deployment is suggested in the discussion section. Then follows the methodology section, including explaining the learning machine-based media analysis that was used to demonstrate the importance of visibility for technology acceptance. Finally, discussion, conclusions, and policy implications are presented. This methodology is rather new and experimental, but its main contribution is highlighting the paradigm shift from human-made media analysis to machine-made analysis with a multidisciplinary approach, and describe its possibilities in technology intelligence, especially in weak-signal detection related to long term R&D strategy decisions.

2.1 Public acceptance of CCS

The viability of CCS, or any other technology, is influenced by economic, regulatory, and technical aspects, but also by public acceptance. Public acceptance of CCS is seen to depend on people's sense of trust in stakeholders and not solely on the properties of the technology itself. (Terwel et al. 2011). The size of the project and local history as well as trust in stakeholders may influence local public acceptance of CCS (Dütschke 2011). Trust in organisations also affects people's perceptions of the magnitude of risk and the benefits as well, impacting their acceptance of CCS (Terwel et al. 2009). Similar logic has been presented, for example, for public acceptance of gene technology (Siegrist 2000) and also for nuclear waste where overwhelming political opposition has been fueled by the public's perception of risks (Slovic et al. 1991).

Education about CCS can also affect public acceptance by highlighting qualities of the technology that the public finds acceptable and thereby reducing fundamental opposition (Itaoka et al. 2004). Public acceptance of different CCS elements— namely plant type, transport, and storage—may, however, be different, as Wallquist et al. (2012) indicate. Pipelines, for example, may result in lower acceptance, whereas storage location can have the least influence (although environmental legislation practically prohibits land storage in Europe), and plant type some influence. Itaoka et al. (2009) indicate that different factors, including risks, effectiveness, responsibilities, and fuel use, have varying impacts on CCS acceptance.

Lay attitudes toward CCS are also seen as relevant, and the lack of public acceptance is seen to potentially reduce the viability of CCS severely (Terwel et al. 2009). In fact, people's acceptance is seen as critical for the widespread deployment of any low-carbon technologies to become viable options for reducing CO_2 emissions (Fleishman et al. 2010). The way CCS might contribute to reducing the impact of global warming is unclear, even to those who believe they have a good understanding (de Best-Waldhober et al. 2009). This is interesting, as many studies indicate that awareness of the necessity of preventing global warming can be crucial to the acceptance of CCS (Itaoka et al. 2009; Tokushige et al. 2007)

Past examples exist for lack of public acceptance being a major hindrance for developing new energy infrastructure cost-effectively, affecting many technologies, including nuclear (Grove-White et al. 2006), CCS (Bradbury et al. 2009), wind farms (Firestone and Kempton 2007), gene technology (Siegrist 2000), nanotechnology (Siegrist et al. 2007a), and many others. Public acceptance in these cases is typically affected by fears of radiation (Kim et al. 2013), CO_2 being released from the ground and causing suffocation (Wallquist et al. 2009), potential noise or threat to animals (Wolsink 2007), and unknown consequences (Zechendorf, 1994; Siegrist et al. 2017b).

Public acceptance is somewhat an unknown factor in developing public policy for CCS technology (Itaoka et al. 2004). Only educating in order to increase public awareness of need for mitigating CO_2 emission would not directly increase the acceptability of CCS (Itaoka et al. 2004), but information may increase support for some aspects of the technology, such as storage options. On the other hand, information on CCS may in some cases result in stronger opposition (Palmgren et al. 2004), particularly against geological storage under the ocean. It is noteworthy that public acceptance depends on information sourced from different actors, especially people's influence on each other, emphasizing trust (Huijts et al. 2007). International examples may also be required to enhance confidence and trust in CCS, as public acceptance is seen as a requirement for market deployment (de Coninck et al. 2009). In fact, high public acceptance is seen as one of the critical factors for widespread deployment of various CCS projects (Zhang and Huising 2017).

Public acceptance is seen as one of the important obstacles for CCS implementation, along with a lack of policy framework, costs, and international regulatory framework, a factor that is seen to potentially have the biggest effect on commercial success (Gough 2008). In some ways, however, public acceptance is viewed among other uncertainties surrounding CCS (Lohwasser and Madlener 2012). Benefit and risk perceptions are seen to influence on the progress of the technology (Wallquist et al. 2010).

Wüstenhagen et al. (2007) describes three types of public acceptance to highlight different aspects of market deployment, namely socio-political acceptance, market acceptance and community acceptance. Bell et al. (2007) note how public acceptance can have multiple dimensions by indicating that the acceptance of generic technology might be very different from that of local projects. Regardless of general acceptance of CCS, 'not in my backyard' (NIMBY) attitudes can appear when facilities are proposed close to one's own communities, yet attitudes about CCS are based on concepts and perceptions, not on actual past events, making the possibilities of comparing NIMBY attitudes to other energy industry developments somewhat limited (Krause et al. 2014).

Although there are many CO_2 storage sites available, the possibility of CO_2 leaking from the storage area has affected public opinion towards the technology. Wallquist et al. (2011) found the NIMBY attitudes to exist towards both CO_2 pipelines and storage sites. Such attitudes persist regardless of techno-economic aspects favouring the large technology market deployment of near-zero CO_2 power production in the medium term (10-20 years). Due to public fears, CCS market deployment in the form of building a commercial-size demonstration plant (for example oxyfuel technology) has been delayed (Santos 2015). The situation has been seen to have strong linkages to public acceptance and as well as to political decision-making.

CCS technologies have been increasingly communicated during their development, starting from the early 2000s (Ashworth et al. 2009). The topic has also attracted, to a lesser extent, attention on social media. Due to the fact that CCS technology is still under development, its commercialisation is dependent on public opinion and on related media communication.

Market deployment includes the actions towards managing organisational resources in the marketplace (Slotegraaf et al. 2003), and deployment is the next step after the R&D activities in the product cycle (Midttun and Gautesen 2007). Various factors (political, technological, financial, etc.) can promote market deployment.

CCS market deployment necessitates achieving effective emission reduction incentives alongside public-private funding for R&D (Gielen et al. 2014). From the technological perspective, the energy mix and ambitious CO_2 reduction targets impact market deployment, whereas should coal be part of the energy mix, CCS is seen as the only technological solution worth deploying (Folke et al. 2011).

Investment costs and CO_2 allowance prices strongly influence the market deployment of coal-fired CCS power plants (Lohwasser and Madlener 2012). Money is an important factor in the market deployment of new energy industry solutions that necessitate private finance (Mathews et al. 2010). Market deployment of new technologies such as CCS requires significant investments and entails some technological risks to demonstrate their viability (Burnham et al. 2013).

Attracting the attention of government and industrial sectors is important for CCS market deployment since incentives, financial support, the regulatory system, and venture capital require widespread participation of government and businesses (Dapeng and Weiwei 2009). Complementary policies and incentives are seen to impact market deployment (Grubler and Riahi 2010). Systemic policy strategy is necessary for market deployment to overcome any technology barriers and manage the risks (Åhman et al. 2013). Different types of policies are potentially needed for supporting low-carbon technologies along with the technology maturity to support the level of market deployment (IEA 2010).

Because it comprises the measures that aim at promoting energy technologies from early research to market deployment, an energy technology policy is needed (Ruester et al. 2014). Initiatives such as the Strategic Energy Technology Plan, the technology pillar of the EU's energy and climate policy adopted by the European Union in 2008, are the first steps toward establishing an energy technology policy for Europe. This type of initiative may eventually result in market deployment of key low-carbon technologies at the European level (Fütterer et al. 2014).

Market deployment is potentially hindered by the commonly understood fact that it typically takes some thirty years for a new technology to materialise and to build the necessary expertise, capacity, and knowledge (Kramer and Haigh 2009). Further, those R&D efforts that focus on technologies with modest potential for mitigating climate change result in market deployment initiatives for technologies to remain fragmented (Grubler and Riahi 2010). In the case of CCS, the time is now critical for the potential market deployment (Maddali et al. 2015). Market deployment takes its time as the extensive number of wells required for global scale deployment of CCS limits the possibilities of deploying CCS on a wide scale in a rapid manner (Maddali et al. 2015).

Public opinion and attitudes are reflected in political decision making, impacting policies, regulations, and even finance. Hence, the realities of CCS market deployment can be affected by the public accepting the technology.

2.2 Research methodology

This study is a first attempt to study media image, public acceptance, and product market deployment by first studying the literature and then comparing the results to findings from empirical analysis through opinion mining with learning machine-based media analysis of a vast number of editorial and social media sources. Therefore, this work is not directly related to one specific field of study; supporting literature is gathered from CI/MI and technology intelligence methods, as well as from corporate decision-making, and is used to describe a possible link from SoMe users to possible effects in company management. The basic research principles have been used in different fields, but are now applied to a single case; in the same way, public acceptance studies have been carried out on other topics using media analysis but with much smaller data sets. Bursher et al. (2015) applied a similar approach with editorial content media framing and sentiment analysis by software. In this study the application of media framing, cluster analysis and statistical methods were considered to be non-applicable. This is due to the comparison of editorial content with social media and to the fact that media frame comparability between two different types of communication is challenging with a large amount of data. Hence, the learning machine-based media analysis is applied in this study to demonstrate the importance of visibility, whether it would be a driver for technology acceptance, namely public acceptance and product market deployment, or not.

The main reasons for choosing the opinion mining approach along with the learning machine-based media analysis method was its applicability to large global data sets (both from editorial content and SoMe), fast data processing, and reduced risk of bias caused by human perceptions and interpretations (Matthes & Kohring 2008). The analysis period and data for this study covers one year, including a major international climate conference, the Paris COP21. Much narrower sentiment analyses have previously been carried out in the field of marketing, yet this study applies the existing elements in a new way. The users of the social web now have a new role as data providers, which seems to provide an excellent platform for analysing public attitudes (Penalver-Martinez et al. 2014). By adopting a media analysis approach and a particular tool, the quantity of media sources to be analysed is drastically increased compared to questionnaires and interviews or traditional media analyses. Merely relying on qualitative methods such as research interviews would entail challenges, compared to a global media coverage study. For example, responses can be difficult to code and answers may vary by participant, while respondents can provide socially acceptable responses, telling what is considered acceptable, to the researcher (Sovacool et al. 2012). The analysis in this study was conducted to clarify the social acceptance status of CCS technology in order to investigate the possible connection to recent challenges in technology market deployment. The analysis findings were synthesised to obtain a clear view of the effect of media image, resulting social acceptance on CCS technology development, and related market deployment. Hence, the research setting in this article is media analysis, where media sentiment is analysed to discover possible implications for public acceptance, political decision-making, and technology market deployment.

The methodology used in this study can be considered a fairly new method in media research, especially in a comparison of global editorial media and global social media. In the past, some attempts have been made to create an automated tool for analysing nuclear power acceptance (Reis et al. 2011), but media sentiment has not been clarified to this extent. This study relies on commercial software to mine the opinions relating to CCS, a similar method to that applied by Bursher et al. (2015). Opinion mining can be seen as a highly active research field consisting of natural language processing, computational linguistics, and text analysis technologies with an aim to get various added-value and informational elements from user opinions (Penalver-Martinez et al. 2014).

The analysis was conducted to clarify the CCS technology's media image. Also, the potential effects on social acceptance of technology and its commercialisation were highlighted by comparing literature to data analysis. Hence, the research setting used in this article is media analysis for one case, which is then compared to different, similar analyses (Nuortimo 2017 a&b)

M-Adaptive software is used as the main tool in the learning machine-based analysis of global editorial and social media (SoMe) sources. In this study, the M-adaptive sources cover 3 million social media platforms globally and 100,000 news outlets in 71 languages in 236 regions (M-Brain 2015). Sentiment analysis was carried out based on a combination of linguistic knowledge and human-aided machine learning, which means that the software suggested classifications to researchers who then provided feedback on correctness. By repeating this process a number of times the system learned to improve its classification of content into sentiment categories (M-Brain 2015). In practice, the sentiment-coding expressions in the text were first recognised and classified automatically. The software matched all relevant CCS-related documents after which the sentiment-focused types were assessed, while the overall compound judgement displayed four options: positive, negative, neutral, and mixed. Data analysis was conducted from 4 December 2014–28 February 2016, by searching 'Carbon Capture Storage' and 'CCS', which included a total of 4496 data points (3380 editorial/1116 SoMe).

According to M-Brain's internal tests, 80 percent of the sentiments are correct on average for a given document when using the M-Adaptive software. Hence, it is possible that the system may make a mistake with any given individual document, due to inherent ambiguity in natural language. Further, it is widely known that humans do not agree 100 percent in similar tests either, due to some individuals not being capable of identifying humour or sarcasm. As is the case for any artificial system, humour, sarcasm and irony are beyond the system's abilities to understand. However, catching the trends in the data becomes more accurate as the number of analysed documents increases, meaning that with large volumes, the overall model qualitatively matches human judgement on the same data.

3. RESULTS OF MACHINE-AIDED MEDIA ANALYSIS OF CCS TECHNOLOGY

The large number of data points enabled the analysis of media sentiment towards CCS. Figure 1 depicts overall sentiments towards CCS in both editorial publications and social media.

The number of hits for CCS (4496) was low compared, for example, to wind power during

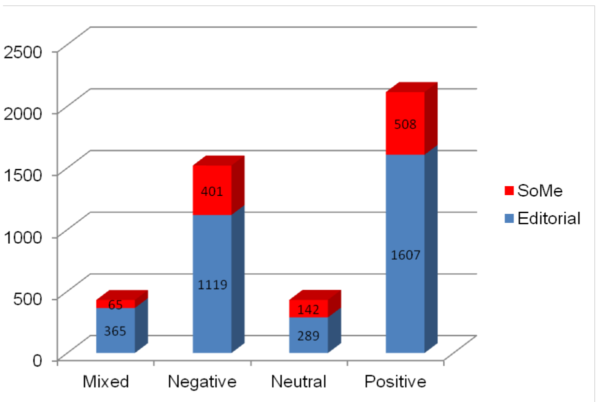


Figure 1 Sentiment analysis of social media and editorial publications.

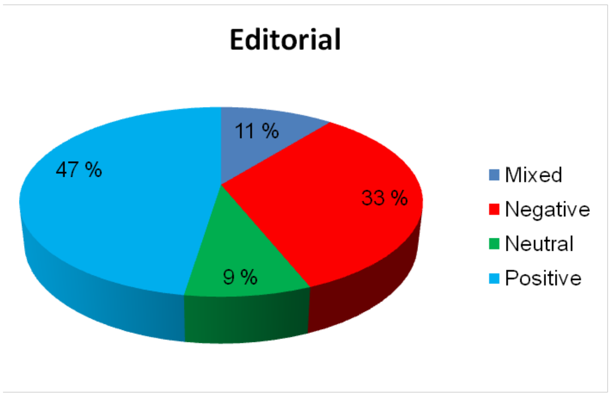


Figure 2 Sentiment analysis of editorial publications.

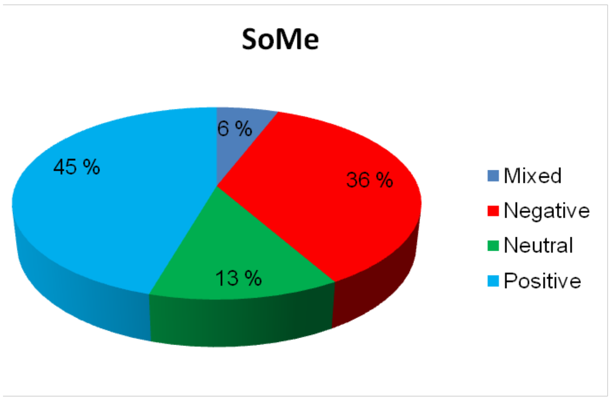


Figure 3 The media sentiment of CCS in social media.

the same period (76,819), indicating relatively low visibility of CCS in the media. The results show that CCS resulted in positive hits mostly in editorial publications but also in social media. Nevertheless, a larger proportion of negative hits in social media indicate lower levels of public technology acceptance. Additionally, the number of SoMe hits is smaller compared to editorial hits, which also indicates less exposure to the general public. Further analysis shows that 33% of hits in the editorial publications were negative and 47% positive, indicating relative technology acceptance among scientists, experts, and journalists. The number of mixed and neutral hits is relatively small, which seems to indicate a consensus towards CCS (Figure 2).

Attitudes in social media appeared somewhat different compared to editorial publications. Figure 3 indicates that public sentiment toward CCS in social media is also mostly positive (45%) with only a minor 2% difference compared to editorial publications. The amount of negative hits was 3% higher than in editorial publications, indicating a bit more negative attitude. In mixed hits the difference was 6-11%, which can be seen as an indication of stricter view expression in social media. However, the 4% more neutral hits seem to indicate that some groups have not yet firmly fixed their attitudes, which can be considered an indication of a need to increase communication efforts in SoMe.

Figure 4 illustrates the social media sentiment of CCS across different media. Dividing the social media sentiment by media type reveals that blog writing has attracted most of the social media attention with over six hundred hits, of which the largest share is positive towards CCS. Also Facebook has been active with over 250, mostly negative, hits. Due to a more visible number of negative hits, the social media effect can be considered quite large when public opinion towards technology is formed.

In Figure 5, media sentiment in selected countries is presented. In Germany, France, and Finland, the sentiment was more positive than in China or Australia, emphasising the need for further communication efforts.

Relevant international events may also influence the appearance of pertinent writings in the media and media sentiment at the time. For example, during the Paris COP negotiations from 30 November to 12 December 2015, a total of 279 hits appeared in the media. The media attention towards CCS was

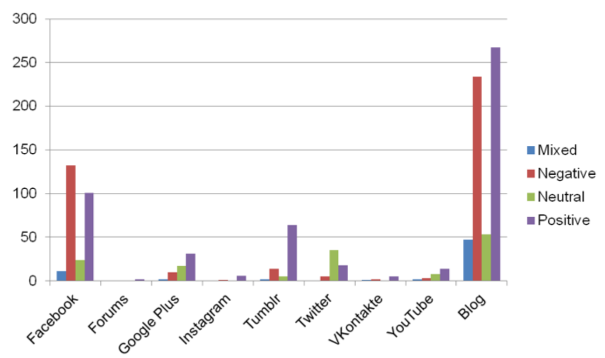


Figure 4 Social media sentiment of CCS across different media.

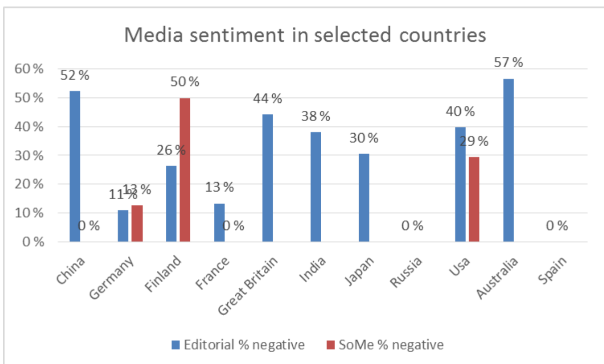


Figure 5 Negative sentiment percentage in selected countries.

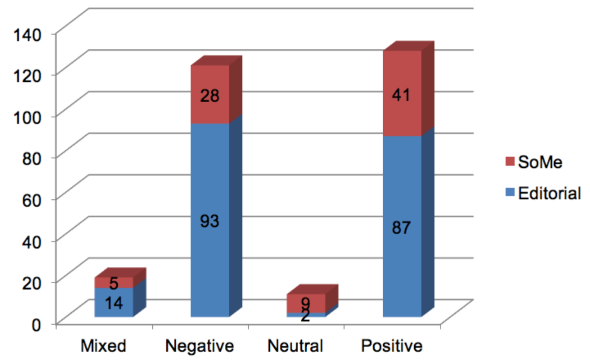


Figure 6 Media hits during Paris COP 30.11-12.12.2015.

approximately doubled during these two weeks compared to an average of 300 hits a month (Figure 6) (calculated as monthly average over 15 months).

Aside from the visibility of CCS being relatively low, it was evident that the editorial hits during the meeting were more negative than usual with 47% negative hits for CCS, while the same for SoMe was only 34%. The normal 15-month averages were 33% and 36%, respectively. The percentages of positive hits during the Paris COP negotiations were 44% and 49%, respectively, while the 15 month averages were 47% and 45%.

4. DISCUSSION

This paper describes the media image of CCS technology, with possible implications especially from SoMe for public acceptance and product market deployment, by synthesising a possible literature-based connection and demonstrating the role of visibility of CCS technology via advanced media analysis. When comparing the literature and empirical findings, the following can be observed. CCS has smaller media exposure with a more positive image. According to some communications theories, large media exposure can have some effect, whether positive or negative; small exposure maybe doesn't affect at all, and small attention is transferred to be negative—if something is unknown, it has more associated risks. Here, this is visible via the number of hits through various media-channels, especially in the editorial/SoMe ratio. When comparing CCS to the case of biomass, CCS also has a positive image with a small number of hits, making the impact smaller. In the case of CCS, one of the main findings is that it is rather unknown, which is the worst case, because people can be afraid of what they don't know. This is evident both from literature as well as from our analysis, therefore partly validating the method used.

In the case of CCS, both communication and corporate stakeholder literature prove beneficial for explaining the phenomenon. For example, traditional stakeholder salience theory does not fully take into account general public attitudes, which can influence corporate decisions both directly and indirectly. In the case of CCS, it is evident that: 1) Literature states that CCS is unknown (Wallquist et al. 2011), which is empirically true due to low numbers of media hits. 2) PR-communication theory implies that if technology is unknown, it can have poor acceptance (McCorkindale et al. 2013). This is evident via the opposition to end storage in different countries and single projects. Also, empirical country by country analysis indicates a high percentage of negative hits in countries with no deployment, such as Australia, and also a high percentage of negative hits in SoMe, such as in Finland. 3) Communication has been intra- and interspecialistic (Ashworth et al. 2009). This follows the funnel model by Bucci et al. (2008). This is empirically visible via the low number of hits, indicating the urgency to increase communication activities to the general public already in the beginning of the product development cycle. 4) Poor media image can possibly have an effect on technology market deployment in the case of CCS. This can be deducted from points 1–3. 5) Means to measure media image have previously been challenging to apply to large global data sets. This study incorporates a new method, opinion mining approach including machine learning, which is tested and found applicable for fast large data-set sentiment analysis.

The total media sentiment relating to CCS was found to be generally positive based on the analysis due to a relatively large number of positive editorial hits, among the rather low media visibility. In the social media, the sentiment seemed to be a bit more negative. For example, Facebook appeared as a platform with active discussions concerning CCS with over 250, mostly negative, hits. The appearance of CCS in various platforms used by the public highlights the role of social media in shaping opinions.

The sentiment also varies by country, as, for example, Germany and France had positive attitudes, whereas Australia had a negative media sentiment, with no deployment of the technology possibly twined with the sentiment. The sentiment can also vary among the type of media, as, for example, in Finland, the editorial content was seen to be more positive than in the social media. The general attitude towards the technology may differ from the local as for example in Germany, it seems that NIMBY is large, regardless of positive general attitudes in both editorial and SoMe content, and as projects have been cancelled due to challenges in finding end-storage sites. Such matters are not directly visible in media analysis and therefore this is a limitation of the utilised methodology. The analysis, however, indicates that general public opinion can be an important factor for public acceptance, and derived from that aspect, also for political decision making. Hence, from the perspective of market deployment, it seems that the more editorial and SoMe content CCS can obtain the better, to counteract the status of being unknown, whereas all possible scientific, technical, marketing and PR communication efforts are important for CCS market deployment, especially those targeted to the general public.

The media sentiment toward a technology can be affected temporarily by relevant international events, such as the global climate negotiations, Paris COP 21, during which the media sentiment seems to be influenced in one way or another. In this case the effect towards CCS by the editorial publications was mostly negative.

Although the needs of CO_2 reduction and the related agreements are of a global nature, technology commercialisation is influenced by regional politics and legislation. It is to be noted that local NIMBY attitudes are not necessarily clearly visible by using the approach in this study. Any discrepancies between media sentiment and the actual project implementation seem to be a clear indication of stronger NIMBY attitudes.

It would seem that one of the main benefits of the study lies in discovering global trends and technology development directions with a larger data set than previous studies, and also trying to establish new methodology for big-data-based media research. Also, this study highlights effectively the differences in channels of communication that may affect public acceptance and perhaps political decision making. The role of SoMe is continuously increasing and presents a challenge for technology developers. It seems that at some level, a speculative negative link from public acceptance, economics, and policies to technology market deployment might exist in the case of CCS.

Another contribution of this study lies in incorporating a method formerly utilised mainly for marketing purposes to study media image and, furthermore, trying to find correlations to public acceptance of CCS, therefore bringing a new angle to related media and social acceptance issues. This is a new approach compared to questionnaire- or interview-based studies with moderate data sets of some hundreds of data points that are used in similar studies (e.g. Heras-Saizarbitoria et al. 2011). When compared to regular qualitative studies, the method has its positives and negatives, but it can be considered an approach that might provide a basis for longitudinal data-series analysis in the future.

As highlighted by Sovacool (2013), quantitative tools can make it difficult to indicate nuances and variance, and they also seldom look for acceptance. However, by utilising this method and comparing editorial content and SoMe, some indication of acceptance appears to have been gained. Hence, it is straightforward that this type of approach would be best, if supplemented with qualitative methods, such as questionnaires. The software sets some limitations, although it still allows the analysis of extensive data sets. The important local media sentiments, such as the NIMBY syndrome (Wolsink 2000), have not been analysed.

In accordance with the results by Heras-Saizarbitoria, et al. (2011), it would seem to be a call for research combining qualitative and quantitative study on the public acceptance issue of CCS technologies. The type of approach involving vast data might be most useful to sight larger trends and could be complimented by qualitative methods, such as questionnaires and interviews. Also further text analysis methods could be applied, such as framing and discourse analysis, but as in this case, the comparability of two large data sets can be challenging. This is due to different types of communication in SoMe, such as hate speech. The changes that take place in the mass media coverage and framing can also affect public acceptance (Heras-Saizarbitoria, et al. 2011). However, this is not so visible when using this type of approach. Also, these types of issues are often emotionally charged, potentially influencing the appearance of the issue, particularly in social media. According to Stieglitz and Dang-Xua (2013), emotionally charged social media messages tend to be repeated more often and more quickly compared to neutral ones. Hence, there is a possibility that media sentiment is influenced by these types of factors.

The managerial implications of this study are related to MI/CI method utilization, and also public acceptance research method development issues. This study highlights the fact that in traditional stakeholder theories, a SoMe participant is not considered so much as a salient stakeholder. However, when combining SoMe users into larger groups, there are possible implications at the corporate level in cases needing both proper political decisions and regulatory environment and policies, as well as long-time R&D activities with also perceived technical and HSE risks. This study tries to find applications of a new method for power plant investment-related media analysis, a learning machine-based sentiment analysis that utilises a very large global data set. Managers working with relevant issues can potentially benefit from the results or the potential of the methodology. The method is applicable to analysing global attitudes, and also their changes, for example, during the time of relevant international events. Furthermore, managers planning power projects or long-term R&D development projects may benefit from understanding the needs for public engagement, and the urgency of social media participation. Figure 7 describes a possible chain from CCS media-image to product market deployment.

This chain starts from public image, which influences people's perceptions of technology. In addition to traditional news media, which can shape public opinion regarding any issue by emphasising certain elements of the broader controversy over others (Shah et al. 2002), social Media (SoMe) presents more direct opinions, often including emotional content (Stieglitz and Dang-Xua 2013). The application of social media is seen to support market intelligence and product development (Berendsen et al. 2015). Media framing in editorial content has the potential to influence public acceptance as attention is focused and placed on a field of meaning (Heras-Saizarbitoria, et al. 2011). Following this reasoning, in PR-communication literature, the rule of effects describes the chain from media exposure via attention, comprehension, motivation, and behavioural trial to sustained behavioural change (McCorkindale et al. 2013). According to the rule of effects, in the rule of halves describing the effect is halved in each step, leaving the percentage from media exposure to sustained behavioural change to 0.78 %, emphasising the need for extensive media exposure. For CCS, one main challenge when the public perception is considered is that in most countries, the public is rather unfamiliar with the technology (Wallquist et al. 2011). This also seems to indicate that communication activities so far have been mostly intra- and interspecialistic, following the funnel model by Bucci et al. (2008), which states that more popular communication is usually done in the commercialisation stage of the product development.

Media image influences public acceptance, and furthermore, public opposition can influence CCS projects directly in the form of local action groups, and indirectly via making the political climate unfavourable for CCS (Wallquist et al. 2011). Recent years have witnessed proliferation of studies on public perceptions of CCS, accompanied by the efforts to translate such knowledge into toolkits for public engagement and communication. At the same time, both literature and toolkits have paid little attention to the organisational dynamics and views of project implementers with regard to public engagement (Breukers et

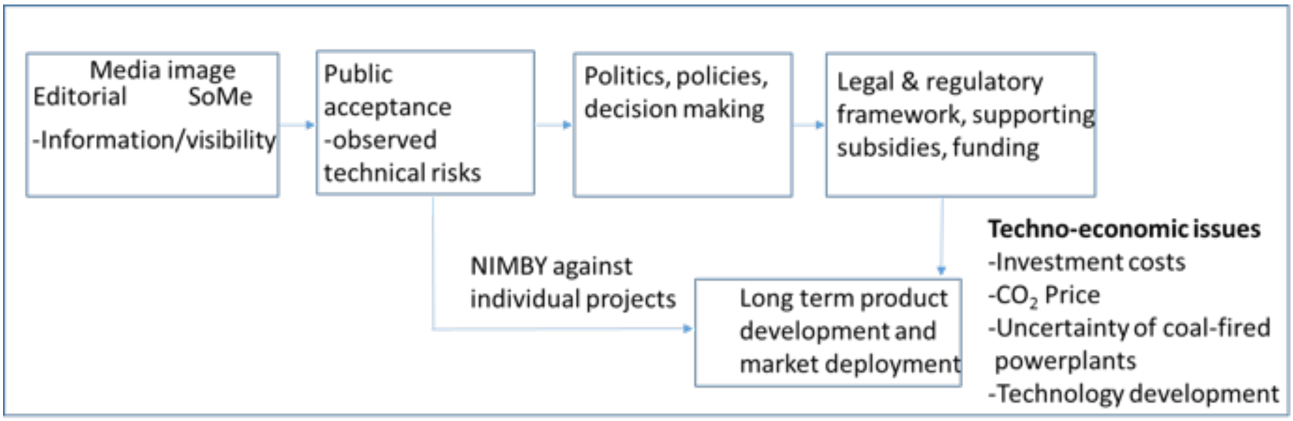


Figure 7 Possible chain from media image to product market deployment/case CCS.

al. 2015). Allowing for improved understanding of the global capacity and applicability of CCS is seen to potentially strengthen the global trust, awareness, and public confidence in CCS technology (de Coninck et al. 2009). For nuclear waste, it was observed that long-term, stable contacts with the local politicians and population are important, but also, as can be seen from the Finnish decision by Parliament, a good contact with the national politicians is necessary. However, there is not necessarily a link between national public acceptance (or lack of it) and political decisions. National decisions, however, require a local acceptance (Le Bars, Y., et al.).

A US-based study found that individually, both CCS and biomass are perceived generally as beneficial for energy development by the news media, though they are not often mentioned in combination, as Feldpausch-Parker et al. (2015) emphasise their value for climate change mitigation and as an alternative to fossil fuels. Earlier examples of failed technology commercialisation have indicated that social acceptance is a decisive factor for technologies, including CCS, while the early adoption of the general public may be essential for technology acceptance (Ashworth et al. 2009).

As a final step from public acceptability to managerial decision-making and technology deployment, a stakeholder salience model (Mitchell et al. 1997) can be considered. The stakeholder salience model introduces three key attributes for stakeholder classification: power, legitimacy, and urgency. The question is: how can one evaluate the groups communicating via SoMe? How can one measure someone's power, legitimacy, or urgency when posting opinions in various discussion forums or on Twitter? Considering development and technology deployment of a single company, these groups have seemingly no power, legitimacy, or urgency and could therefore be considered traditionally to be non-stakeholders in the decision making and would be perceived as having no salience by the firm's managers. However, reflecting on Figure 7, in the case of CCS product market deployment, one pathway for this is suggested.

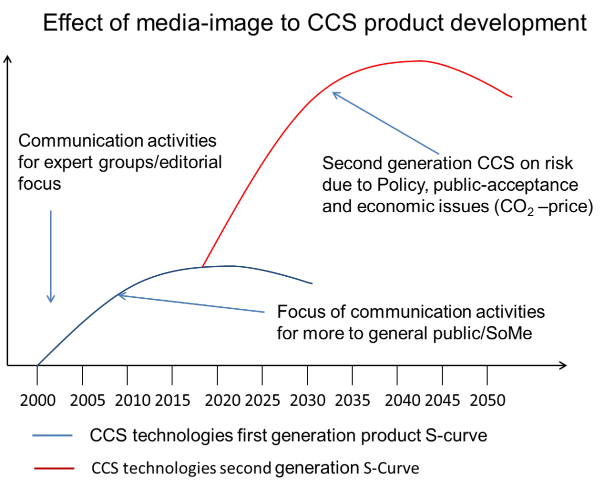


Figure 8 Stakeholder adoption in CCS product development.

Furthermore, Figure 8 is synthesised, suggesting that earlier stakeholder adoption would benefit from CCS market deployment. The findings from media study support this hypothesis via implicating negative attitudes toward the technology, especially in SoMe, and low levels of hits in general, implying unknown technology. The figure illustrates how CCS-technology development would have potentially benefited from the earlier stakeholder adaptation. Furthermore, due to lack of public acceptance, second generation CCS-technology, development is under risk.

Some of the managerial implications of this paper are also related to the R&D decision-making process and the social media influence. This study indicates that investments in CCS technology may not be favourable due to uncertainties in public acceptance. It was clearly visible that the amount of media attention was not large enough to fully support product commercialisation. The utilised artificial learning machine-based analysis tool may prove beneficial when evaluating social acceptance issues affecting long-term R&D investments. Hence, as a practical implication, this study emphasises the need for more versatile analysis of factors affecting long-term R&D investments with strong public involvement both directly and via political decision-making.

The limitations of this study include the analysed media sentiment being limited to those classifications possible with the used keywords and also to the English language. Using other keywords, or not including some topics, might provide slightly different results. In addition, framing, cluster analysis, and statistical methods were found difficult to apply as the comparability between editorial content and SoMe could have been lost. In addition, although statistical techniques are widely used among communications scholars to identify news frames, they are criticised for not being able to do so in a conceptually valid manner (Carragee & Roefs 2004). This also brings a challenge to further research.

The utilised method may entail some uncertainties that require further studies. Results correlate to literature, so that based on the analysis, CCS is unknown and also has more positive sentiment. Also, the methods that were used for CCS product life-cycle estimation are not based on calculated figures and are only directional. In addition to addressing the limitations of this study, relevant future research could relate to developing the machine/artificial intelligence-based methods further.

5. CONCLUSIONS AND POLICY IMPLICATIONS

New AI, computational linguistics and machine learning methods can be utilized for weak signal detection in CI/MI and strategic planning functions of a company. Public acceptance appears as a clearly essential part of the energy market products' market deployment, an issue that should be addressed during the early stages of a product life-cycle. The overall visibility of a technology is important, while if public acceptance is ignored, it can cause delayed or abandoned market deployment of long-term energy production technologies, accompanied by techno-economic issues. This paper has twofold implications. Firstly, it studies CCS media image with a new type of method, public acceptance, and product market deployment based on literature. Secondly, it highlights the importance of visibility and studies possibilities for closing the gap between the rhetoric and technical progress inherent to CCS, which is critically important to global climate mitigation efforts. Developing strong international cooperation to demonstrate CCS with global coordination, transparency, cost-sharing, and communication as guiding principles would facilitate efficient and cost-effective collaborative global learning about CCS. Founded on the learning machine-based media analysis, it appears that the popular type of communication might have been beneficial to start to a larger extent during the early stages of CCS product development.

As a policy implication, the media image of technologies, possibly affecting larger audience groups' public acceptance, can be studied by means of learning machine-based analysis. This type of analysis indicates the majority of attitudes in both editorial publication and social media. Learning machine-based analysis provides a fast way for policy makers to get information on the general public sentiment. The media image of CCS was found to be mainly positive—however, small and unknown, implying a need to push towards regulations to provide some common ground to commercialise CCS technologies. However, the visibility of CCS is currently lacking. Policies favouring CCS could be created as an implication of positive media image; however, not in my back yard (NIMBY) attitudes need to be assessed and addressed locally.

As a future field of study, the further evaluation of adoptability of this type of opinion mining approach to weak signal detection of BI/MI activities is one topic to consider. Additionally, opinion mining method development and the application itself would be an interesting field of study.

6. REFERENCES

Abrahams, A. S., Jiao, J., Fan, W., Wang, G. A., and Zhang, Z. (2013). What's buzzing in the blizzard of buzz? automotive component isolation in social media postings. *Decision Support Systems*, 55 (4): 871.

Åhman, M., Nikoleris, A., and Wyns, T. (2013). Decarbonizing industry: emerging roadmaps point to major need for financing radical innovation. *Carbon Management*, 4 (1): 5-7.

Alphen, K., Voorst tot Voorst, Q., Hekkert, M., and Smits, R. (2007). Societal acceptance of carbon capture and storage technologies. *Energy Policy*, 35 (8): 4368–4380.

Ashworth, P., Boughen, N., Mayhew, M., and Millar, F. (2009). An integrated roadmap of communication activities around carbon capture and storage in Australia and beyond. *Energy Procedia.* 1 (1): 4749-4756.

Le Bars, Y., Murray, C., Ormai, P., Seppälä, T., and Bøhmer, N. Summary of the Panel Discussion Concluding Session IV on Public Involvement and Acceptance Panel IV: '*Can NIMBY be overcome?'*

Bell, D., Gray, T., and Haggett, C. (2005). The 'Social Gap' in Wind Farm Siting Decisions: Explanations and Policy Responses. *Environmental Politics*, 14 (4): 460-477.

Berendsen, G., Middel, R., Pieters, I., Angard, F., and Hillerström, F. (2015). Social media within sustainable product development: an exploratory multiple case study on the perception of social media usability in the new product development process. *Int. J.*

Technology Intelligence and Planning, 10 (3/4): 273–293.

Bradbury, J., Ray, I., Peterson, T., Wade, S., Wong-Parodi, G., and Feldpausch, A. (2009). The role of social factors in shaping public perceptions of CCS: Results of multi-state focus group interviews in the U.S. *Energy Procedia*, 1 (1): 4665–4672.

Brouwer, A. S., van den Broek, M., Seebregts, A., and Faaij, A. (2015). Operational flexibility and economics of power plants in future low-carbon power systems. *Applied Energy*, (156): 107-128.

Bucchi, M. (2008). Of deficits, deviations and dialogues: Theories of public communication of science. Featured in: M. Bucchi and B. Trench (Ed.) *Handbook of public communication of science and technology,* New York: Routledge International Handbooks, 57-76.

Budd, J.M. (2007). Information, analysis, and ideology: A case study of science and the public interest. *Journal of the American Society for Information Science and Technology*, 58 (14): 2366-2371.

Budescu, D. V., Broomell, S., and Por, H.-H. (2009). Improving Communication of Uncertainty in the Reports of the Intergovernmental Panel on Climate Change. *Psychological Science 2*0 (3): 299-308.

Burnham, J., Debande, O., Jones, O., Mihai, C., Moore, J. and Temperton, I. (2013). Report on Innovative Financial Instruments for the Implementation of the SET Plan, First-of-a-kind project. *JRC Scientific and Policy Reports*, Ref: EUR 26058, OPOCE LD-NA-26058-EN-N, https://ec.europa.eu/jrc/sites/default/files/ldna26058enn_002.pdf, accessed 23 May 2017.

Burscher, B., Vliegenthart, R., and de Vreese, C. (2015). Frames Beyond Words: Applying Cluster and Sentiment Analysis to News Coverage of the Nuclear Power Issue. *Social Science Computer Review*, 1-16

Breukers, S. and Upham, P. (2015). Organisational aspects of public engagement in European energy infrastructure planning: the case of early-stage CCS projects. *Journal of Environmental Planning & Management*, 58 (2): 252-269.

Cambria, E., Schuller, B., Xia, Y., and Havasi, C. (2013). New avenues in opinion mining and sentiment analysis. *IEEE Intelligent Systems* 28 (2): 15-21.

Carragee, K. M. and Roefs, W. (2004). The neglect of power in recent framing research. *Journal of Communication*, 54: 214–233.

Chen, H. (2010). Business and market intelligence 2.0. IEEE Intelligent Systems 25(1): 68-71.

Chen, H., Chiang, R. H., and Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly*, 1165-1188.

Dapeng, L. and Weiwei, W. (2009). Barriers and incentives of CCS deployment in China: Results from semi-structured interviews, *Energy Policy*, 37 (6): 2421–2432.

de Best-Waldhober, M., Daamen, D, and Faaij, A. (2009). Informed and uninformed public opinions on CO_2 capture and storage technologies in the Netherlands, *International Journal of Greenhouse Gas Control*, 3 (3): 322–332.

de Coninck, H., Stephens, J. C. , and Metz, B. (2009). Global learning on carbon capture and storage: A call for strong international cooperation on CCS demonstration. *Energy Policy*, 37 (6): 2161-2165.

Dessler, A. E. and Parson, E. A. (2006). *The Science and Politics of Global Climate Change: A Guide to the Debate.* Cambridge: Cambridge University Press

Dütschke, E. (2011). What drives local public acceptance – comparing two cases from Germany, *Energy Procedia*, 4: 6234–6240.

Du Toit, A. S. (2015). Competitive intelligence research: An investigation of trends in the literature. *Journal of Intelligence Studies in Business*, 5(2).

Druckman, J. (2004). Priming the vote: Campaign effects in a U.S. Senate Election. *Political Psychology* 25 (4): 577-594.

Feldpausch-Parker, A., Burnham, M., Melnik, M., Callaghan, M. L., and Selfa, T. (2015). News Media Analysis of Carbon Capture and Storage and Biomass: Perceptions and Possibilities. *Energies*, 8 (4): 3058-3074.

Firestone, J. and Kempton, W. (2007). Public opinion about large offshore wind power: Underlying factors, *Energy Policy*, 35 (3): 1584–1598.

Fleishman, L. A., De Bruin, W. B., and Morgan, M. G. (2010). Informed Public Preferences for Electricity Portfolios with CCS and Other

Low-Carbon Technologies, *Risk Analysis*, 30 (9): 1399–1410.

Folke, C., Radgen, P., Rode, H., Irons, R., Schaaf, H., Read, A., Möller, B. F., Schoenmakers, H., Imber, P., and Peter, K. (2011). E.ON's current CCS activities, *Energy Procedia*, 4: 6091-6098.

Fütterer, M. A., Carlsson, J., de Groot, S., Deffrennes, M., and Bredimas, A. (2014). European energy policy and the potential impact of HTR and nuclear cogeneration, *Nuclear Engineering and Design* 27: 73–78.

Gerbner, G. and Gross, L. (1976). Living with the television: The violence profile. *Journal of Communication* 26 (2): 172-194.

Gielen, D., Podkanski, J., and Unander, F. (2004). Prospects for CO_2 capture and storage Organization for Economic Co-operation and Economic Development, (OECD)/International Energy Agency (IEA), Paris, France, p.252.

Godbole, N., Srinivasaiah, M., and Skiena, S. (2007). Large-scale sentiment analysis for news and blogs. ICWSM, 7 (21): 219-222.

Gough, C. (2008). State of the art in carbon dioxide capture and storage in the UK: An experts' review. *International Journal of Greenhouse Gas Control* 2 (1): 155–168.

Greenberg, M. and Truelove, H. B. (2011). Energy choices and risk beliefs: Is it just global warming and fear of a nuclear power plant accident? *Risk Analysis*, 31 (5): 819-831.

Grove-White, R., Kearnes, M., Macnaghten, P., and Wynne, B. (2006). Nuclear futures: Assessing public attitudes to new nuclear power, *Political Quarterly*, 77 (2): 238–246.

Grubler, A. and Riahi, K. (2010). "Do governments have the right mix in their energy R&D portfolios?", *Carbon Management*, 1 (1): 79-87.

Gupta, N., Fischer, A. R. H., and Frewer, L. J. (2012). Socio-psychological determinants of public acceptance of technologies: A review. *Public Understanding of Science* 21 (7): 782-795.

Hammond, G. P., Akwe, S. S. O., and Williams, S. (2011). Techno-economic appraisal of fossil-fuelled power generation systems with carbon dioxide capture and storage. *Energy*. 36 (2): 975-984.

Heras-Saizarbitoria, I., Cilleruelo, E., and Zamanillo, I. (2011). Public acceptance of renewables and the media: An analysis of the spanish PV solar experience. *Renewable and Sustainable Energy Reviews*, 15 (9): 4685-4696.

Huijts, N. M. A., Midden, C. K. H., and Meijnders, A. L. (2007). Social acceptance of carbon dioxide storage, *Energy Policy*, 35 (5): pp.2780–2789.

Huijts, N. M. A., Molin, E. J. E., and Steg, L. (2012). Psychological factors influencing sustainable energy technology acceptance: A review-based comprehensive framework. *Renewable and Sustainable Energy Reviews,* 16 (1) 525– 531.

IEA. (2010). Energy technology perspectives: Scenarios & Strategies to 2050, International Energy Agency, Paris, <online> https://www.iea.org/publications/freepublications/publication/etp2010.pdf accessed 23 May 2017.

Itaoka, K., Saito, A., and Akai, M. (2004). Public acceptance of CO_2-capture and storage technology: a survey of public opinion to explore influential factors, the Seventh International Conference on Greenhouse Gas Technologies, Vancouver, 5 September 2004.

Kim, Y., Kim, M., and Kim, W. (2013). Effect of the Fukushima nuclear disaster on global public acceptance of nuclear energy. *Energy Policy*, 61: 822–828.

Kramer, G. J. and Haigh, M. (2009). No quick switch to low-carbon energy. *Nature,* 462: 568-569

Krause, R. M., Carley, S. R., Warren, D. C., Rupp, J. A., and Graham, J. D. (2014). 'Not in (or Under) My Backyard': Geographic Proximity and Public Acceptance of Carbon Capture and Storage Facilities. *Risk Analysis*, 34 (3): 529–540.

Lohwasser, R. and Madlener, R. (2012). Economics of CCS for coal plants: Impact of investment costs and efficiency on market diffusion in Europe. *Energy Economics*, 34 (3): 850–863.

Li, Y. and Li, T. (2013). Deriving market intelligence from microblogs doi:https://doi.org/10.1016/j.dss.2013.01.023

Liu, B. and Zhang, L. (2012). A survey of opinion mining and sentiment analysis. *Mining text data*. Springer. 415-463.

Mathews, J. A., Kidney, S., Mallon, K., and Hughes, M. (2010). Mobilizing private finance to drive an energy industrial revolution. *Energy Policy*, 38 (7): 3263-3265.

Maddali, V., Tularam, G. A., and Glynn, P. (2015). Economic and Time-Sensitive Issues Surrounding CCS: A Policy Analysis. *Environmental Science & Technology*, 49 (15): 8959-8968.

Midttun, A. and Gautesen, K. (2007). Feed in or certificates, competition or complementarity? Combining a static efficiency and a dynamic innovation perspective on the greening of the energy industry, *Energy Policy*. 35 (3): 1419-1422.

Mitchell, R. K., Agle, B. R., and Wood, D. J. (1997). Toward a theory of stakeholder identification and salience: Defining the principle of who and what really counts. *Academy of Management Review* 22 (4): 853-886.

McCorkindale, T., DiStaso, M. W., and Carroll, C. (2013). The power of social media and its influence on corporate reputation. The Handbook of Communication and Corporate Reputation, 497-512.

Matthes, J. and Kohring, M. (2008). The content analysis of media frames: Toward improving reliability and validity. *Journal of Communication*, (58): 258–279.

M-Brain. Corporate communications. Received 11/2015.

Miller, J. and Krosnick, J. (2000). News media impact on the ingredients of presidential evaluations: Politically knowledgeable citizens are guided by a trusted source. *American Journal of Political Science,* 44 (2): 295-309.

Nasukawa, T. and Yi, J. (2003). Sentiment analysis: Capturing favorability using natural language processing. Proceedings of the 2nd International Conference on Knowledge Capture, 70-77.

Nuortimo, K. (2012). Essays on exploring alternatives in CO_2-neutral power generation product development. Licentiate thesis. University of Oulu.

Nuortimo, Kalle & Harkonen, Janne & Karvonen, Erkki. (2017). Exploring the global media image of solar power. Renewable and Sustainable Energy Reviews. 81. 2806-2811. 10.1016/j.rser.2017.06.086.

Nuortimo, K., Härkönen, J., and Karvonen, E. (2017b). Exploring the social acceptance of biomass power, *Interdisciplinary Environmental Review*, 18 (1): 14–27.

Oltraa, C., Salaa, R., Solàa, R., Di Massob, M., and Rowec, G. (2010). Lay perceptions of carbon capture and storage technology. *International Journal of Greenhouse Gas Control*, 4 (4): 698–706.

Palmgren, C. R., Morgan, M. G., de Bruin, W. B., and Keith, D. W. (2004). Initial Public Perceptions of Deep Geological and Oceanic Disposal of Carbon Dioxide, *Environmental Science & Technology*, 38 (24): 6441–6450.

Peñalver-Martinez, I., Garcia-Sanchez, F., Valencia-Garcia, R., Rodríguez-García, M. Á., Moreno, V., Fraga, A., and Sánchez-Cervantes, J. L. (2014). Feature-based opinion mining through ontologies. *Expert Systems with Applications*, 41 (13): 5995-6008. doi:10.1016/j.eswa.2014.03.022

Pietilä, V. (1997). Joukkoviestintätutkimuksen valtateillä. Tampere: Vastapaino. In Finnish.

Rainer, D. (2008). *A Looming Rhetorical Gap: A Survey of Public Communications Activities for Carbon Dioxide Capture and Storage Technologies*. Judge Business School, University of Cambridge.

Reis, T., Barroso, A. C., and Imakuma, K. (2014). Monitoring and analysis of nuclear acceptance by information retrieval and opinion extraction on the internet. Available from Thiago Reis, retrieved on 21 June 2016.

Ruester, S., Schwenen, S., Finger, M., and Glachant, J.-M. (2014).A post-2020 EU energy technology policy: Revisiting the strategic energy technology plan, *Energy Policy*, 66: 209–217.

Santos, S. (2015). Oxy-CFB Combustion Technology, Its potential role in CO_2-mitigation; O_2GEN workshop, 18 June 2015, Turku, Finland.

Scheufele, D. (1999). Framing as a Theory of Media Effects. Journal of Communication, 49 (1): 103-122.

Shah, D. V., Watts, M. D., Domke, D., and Fan, D. P. (2002). News framing and cueing of issue regimes: Explaining Clinton's public approval in spite of scandal. *Public Opinion Quarterly*, (66): 339–370.

Slotegraaf, R. J., Moorman, C., and Inman, J. J. (2003).The Role of Firm Resources in Returns to Market Deployment, *Journal of Marketing Research*, 40 (3): 295-309.

Song, M. and Thieme, J. (2009). The role of suppliers in market intelligence gathering for radical and incremental innovation. *Journal of Product Innovation Management* 26(1), 43-57.

Sovacool, B. and Ratan, P. L. (2012). Conceptualizing the acceptance of wind and solar electricity. *Renewable and Sustainable Energy Reviews*, 16 (7): 5268–5279.

Siegrist, M. (2000).The influence of trust and perceptions of risks and benefits on the acceptance of gene technology, *Risk Analysis* 20 (2): 195–203.

Siegrist, M., Cousin, M.-E., Kastenholz, H., and Wiek, A. (2007a). Public acceptance of nanotechnology foods and food packaging: The influence of affect and trust, *Appetite*, 49 (2): 459–466.

Siegrist, M., Keller, C., Kastenholz, H., Frey, S., and Wiek, A. (2007b). Laypeople's and Experts' Perception of Nanotechnology Hazards, *Risk Analysis* 27 (1): 59–69.

Slovic, P., Flynn, J., and Layman, M. (1991). Perceived Risk, Trust, and the Politics of Nuclear Waste. *Science* 254: 1603–1608.

Stieglitz, S. and Linh D.-X. (2013). Emotions and Information Diffusion in Social Media—Sentiment of Microblogs and Sharing Behavior. *Journal of Management Information Systems* 29 (4): 217-248

Teo, T. S. and Choo, W. Y. (2001). Assessing the impact of using the internet for competitive intelligence. *Information & Management*, 39(1), 67-83.

Ter Mors, E., Weenig, M. W. H., Ellemers, N., and Daamen, D. D. L. (2010). Effective communication about complex environmental issues: Perceived quality of information about carbon dioxide capture and storage (CCS) depends on stakeholder collaboration. *Journal of Environmental Psychology*, 30 (4): 347-357.

Terwel, B. W., Harinck, F., Ellemers, N., and Daamen, D. D .L. (2009). Competence-Based and Integrity-Based Trust as Predictors of Acceptance of Carbon Dioxide Capture and Storage (CCS). *Risk Analysis* 29 (8), 1129–1140.

Terwel, B. W., Harinck, F., Ellemers, N., and Daamen, D. D. L. (2011). Going beyond the properties of CO_2 capture and storage (CCS) technology: How trust in stakeholders affects public acceptance of CCS. *International Journal of Greenhouse Gas Control*, 5 (2): 181–188.

Tokushige, K., Akimoto, K., and Tomoda, T. (2007). Public perceptions on the acceptance of geological storage of carbon dioxide and information influencing the acceptance, *International Journal of Greenhouse Gas Control*, 1 (1): 101–112.

Valkenburg, P. and Valkenburg, W. (2016). Media Effects: Theory and Research.. 67: 315–338.

Wallquist, L., Visschers, V. H. M., and Siegrist, M. (2009). Lay concepts on CCS deployment in Switzerland based on qualitative interviews. *International Journal of Greenhouse Gas Control*, 3 (5): 652–657.

Wallquist, L., Visschers, V. H. M., and Siegrist, M. (2010). Impact of Knowledge and Misconceptions on Benefit and Risk Perception of CCS. *Environmental Science & Technology*, 44 (17): 6557–6562.

Wallquist, L., Seigo, S. L., Visschers, V. H., and Siegrist, M. (2012). Public acceptance of CCS system elements: A conjoint measurement. *International Journal of Greenhouse Gas Control*, (6): 77-83.

Wennersten, R., Sun, Q., and Li, H. (2015). The future potential for Carbon Capture and Storage in climate change mitigation – an overview from perspectives of technology, economy and risk. *Journal of Cleaner Production*, 103 (15): 724-736.

Wolsink, M. (2000). Wind power and the NIMBY-myth: Institutional capacity and the limited significance of public support. *Renewable Energy* 21 (1): 49-64.

Wolsink, M. (2007). Wind power implementation: The nature of public attitudes: Equity and fairness instead of 'backyard motives'. *Renewable and Sustainable Energy Reviews*, 11 (6): 1188–1207.

Wood, E. (2001). Marketing information systems in tourism and hospitality small- and medium-sized enterprises: A study of internet use for market intelligence. *International Journal of Tourism Research*, 3(4): 283-299.

Wüstenhagen, R., Wolsink, M., and Bürera, M. J. (2007). Social acceptance of renewable energy innovation: An introduction to the concept. *Energy Policy*, 35 (5): 2683–2691.

Xu, K., Liao, S. S., Li, J., and Song, Y. (2011). Mining comparative opinions from customer reviews for competitive intelligence. *Decision Support Systems*, 50(4): 743-754.

Zhang, Z. and Huising, D. (2017). Carbon dioxide storage schemes: Technology, assessment and deployment. *Journal of Cleaner Production*, 142, Part 2: 1055–1064.

Zechendorf, B. (1994). What the Public Thinks about Biotechnology. *Bio/Technology*, 12: 870–875.

JISIB

Journal of Intelligence Studies in Business
Vol. 8, No. 2 (2018) pp. 23-42
Open Access: Freely available at: https://ojs.hh.se/


Stevan Dedijer
1911-2004

Social business intelligence: Review and research directions

Helena Gioti[a], Stavros T. Ponis[b*] and Nikolaos Panayiotou[b]

[a] *Hellenic Open University, Greece*
[b] *School of Mechanical Engineering, Section of Industrial Management and Operations Research, National Technical University Athens, Greece*

Corresponding author (*): staponis@central.ntua.gr

Received 14 June 2018 Accepted 20 August 2018

ABSTRACT Social business intelligence (SBI) is a rather novel discipline, emerged in the academic and business literature as a result of the convergence of two distinct research domains: business intelligence (BI) and social media. Traditional BI scientists and practitioners, after an inevitable initial shock, are currently discovering and acknowledge the potential of user generated content (UGD) published in social media as an invaluable and inexhaustible source of information capable of supporting a wide range of business activities. The confluence of these two emerging domains is already producing new added value organizational processes and enhanced business capabilities utilized by companies all over the world to effectively harness social media data and analyze them in order to produce added value information such as customer profiles and demographics, search habits, and social behaviors. Currently the SBI domain is largely uncharted, characterized by controversial definitions of terms and concepts, fragmented and isolated research efforts, obstacles created by proprietary data, systems and technologies that are not mature yet. This paper aspires to be one of the few -to our knowledge- contemporary efforts to explore the SBI scientific field, clarify definitions and concepts, structure the documented research efforts in the area and finally formulate an agenda of future research based on the identification of current research shortcomings and limitations.

KEYWORDS Big data, business intelligence, review, social business intelligence, social media

1. INTRODUCTION

In the last decade, business intelligence (BI) has proved, beyond any doubt, that it is a rapidly expanding domain in both research and business terms with the number of BI related scientific publications and organizations embracing BI methodologies, techniques, tools and platforms rapidly increasing year by year. This remarkable growth is directly connected with the abundance of customer/user data as a result of increased bandwidth, technological advancements in information systems and mobile applications and the explosion of user generated content mostly materialized by social media and other Web 2.0 platforms. Nowadays, social media and BI are converging faster than ever before. The confluence of these two emerging domains is already producing new added value organizational processes and enhanced business capabilities utilized by companies all over the world to effectively harness social media data and analyze them in order to produce added value information such as customer profiles and demographics, search habits, and social behaviors.

This point of convergence is exactly the scientific area where this paper sets its focus and research efforts, i.e. social business intelligence (SBI), a very new concept trying to capture this transformation of BI systems in the era of Big Data and amidst the social media revolution. This paper constitutes the third -to our knowledge- effort to explore the SBI scientific field, clarify definitions and concepts, structure the documented research efforts in the area and finally formulate an agenda of future research based on the identification of current research shortcomings and limitations.

In doing so, this paper follows a structured literature review approach utilizing data from one of the most established academic databases in the world, i.e. Elsevier's SCOPUS, and imposing a 'search and filter' process based on a carefully selected set of inclusion and exclusion criteria described in detail in the next section. The collected papers were studied thoroughly with the objective to initially eliminate duplicates and critically exclude papers dealing with SBI superficially, fragmentally or not at all. At the end of the literature scrutiny process, 83 papers were selected for further in-depth, full-text examination with the objective to provide the reader with an overview of the main themes and trends covered by the relevant SBI literature. The review process imposed on the 83 papers included in the final review sample produced several interesting findings regarding the current structure of the domain and the necessary prioritization of the research activities for the future.

The remainder of this paper is structured as follows. The next section provides a brief theoretical background of the two domains under study, i.e. BI and social media. It aims to provide the necessary information for understanding the importance of Big Data for BI and the potential impact and transformative nature that social media have on existing BI research and practice with a special focus on User Generated Content (UGC) and trends related to specific social media platforms. In Section 3, the methodological approach to the review and the results of the selection process are presented, followed by the review of the selected papers and the synthesis and taxonomization of the identified research efforts in Section 4. Finally, Section 5 attempts a critical discussion of the review findings in Section 4 and concludes with a proposed SBI domain research taxonomy and a suggested list of priorities and directions for future research.

2. BACKGROUND

Business intelligence (BI) is an "umbrella" term including a wide range of processes, applications and technologies through which various data sources can be gathered, stored, accessed, and analysed in order to gain meaningful information crucial for decision-making (Olszak, 2016). The term, although growing in popularity recently, was first introduced more than seven decades ago to describe "*an intelligence system utilising data-processing machines for auto-abstracting, auto-encoding and profiling of action points in an organisation*" (Luhn, 1958).

However, only recently it turned to a prevailing field for academics and practitioners and a leading commercial concern for most business organisations. According to Chen et al. (2012), there are several reasons explaining this incremented popularity. On the one hand, there is a great opportunity from the rapid expansion of readily available web data sources and on the other, BI tools are becoming more sophisticated, easier to use and find applications in many business processes. Meanwhile, intensive competition and global economic pressures set the success barrier too high, leading companies to a continuous fight for improvement, better quality of service (QoS) and more productive operations.

Chaudhuri et al. (2011) underline the declining cost of data storage and acquisition as an additional reason for the extensive proliferation of BI systems. The same applies to hardware, which is becoming more technologically advanced and less expensive, allowing for more powerful architecture of data warehouses.

The implementation of BI provides modern organisations, even SMEs (Ponis et al., 2013), with the ability to achieve timely and quality decision-making, which constitutes a crucial prerequisite to build a stable competitive advantage. Upon the effective aggregation of "intelligent" data regarding the internal and external business environment, executives are able to take proactive actions preparing their firms for future economic trends and conditions. According to Ranjan (2009), BI is like a "crystal ball" in the hands of managers, revealing the best course of action depending on five major parameters: the company's position in relation to its rivals; the overall strengths and weaknesses of the company; current and future market demographic and economic trends; social, political and regulatory environment; competitions' decisions and strategy and finally, customer preferences and purchasing patterns.

Beyond any shadow of doubt, the business landscape in the era of a fast paced and intensively competitive environment is dominated by the struggle to proactively respond to changes, satisfy the increasingly demanding customer needs and timely decision making on the best courses to action. BI and sophisticated analytics provide contemporary enterprises with the tools, methods and corporate mentality required to survive the hard business arena and maintain profitable relationships with the whole value chain surrounding their activities.

The concept of participation, on which Web 2.0 is based, has also great economic implications and opens up significant new potentials for enterprises (Tziralis et al., 2009). In this very demanding and fiercely competitive environment, businesses have found a powerful ally in the face of social media applications and their fast-paced advancement and prevalence in the business and internet ecosystem. Social media are online platforms through which users can communicate, share content and connect with each other. Since their first appearance in the early 2000s, social media are constantly increasing in numbers, types and popularity. According to the academic literature, social media constitute a reasonable aftereffect of Web 2.0, an argument that is summarised in Kaplan and Haenlein's (2010) definition: “Social media is a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, which allows the creation and exchange of user-generated content”. However, what clearly distinguishes social media from other Web 2.0 applications is the element of social connectivity on a personal level. Within such sites, users pre-select their connections and own privacy control over the content they share (Heijnen et al., 2013).

When it comes to classification, there is no systematic way in which social media can be categorised. Indicatively, Diamantopoulou et al. (2010), propose a rational social media segmentation based on their major activity (i.e. communication, collaboration, share, rate and opinions’ expression) and purpose of use (leisure, work/business, democratic engagement). Kaplan and Haenlein (2010), however, suggest a matrix categorisation consisting of two social media dimensions, namely; self-presentation/ self-disclosure and social presence/media richness.

The outburst of social media and their increasing popularity has led to an era of fast and immense internet data generation. Consequently, the notion of social media analytics and its utilisation in BI systems has become a dominant trend in the entrepreneurial world, due to its huge potential in added-value applications (Fan et al., 2015). In the next sections an attempt to structure the current research domain on the intersection of these two disciplines is made, following a systematic literature review approach.

3. RESEARCH METHOD

According to Hart (1998) a literature review is an objective, thorough summary and critical analysis of the relevant available research literature on the topic being studied. A review of prior and relative literature of a scientific area is an essential feature of academic progress and theory development, since it creates a solid foundation for understanding the current research status quo, while at the same time highlights underdeveloped or unexplored areas as candidates for future research. The literature review should contain processed information from all available sources, be unbiased to the highest possible extent, be free from jargon terminology and supported by a well-defined and consistent search and selection strategy (Hart, 1998).

This review examines literature contributions directly addressing SBI, i.e. the use of social media for BI purposes between 2006 and 2016. Expanding the search before 2006 was deemed unproductive since the advent of social media in its current form is connected with the launch of Facebook in 2004. Previous efforts, like Friendster and MySpace, are not taken into consideration in this study, since they never managed to establish their social media presence and were either defunct (Friendster) or forced to pivot their offerings (MySpace).

In this paper, we utilize a systematic literature review (SLR) approach, which is a trustworthy, rigorous and auditable methodology for evaluating and interpreting all available research relevant to a particular research question, topic, area or phenomenon of interest (Keele, 2007). The selection process was straightforward. Initially, it was decided that the SCOPUS academic database was adequate in order to provide this study with a representative list of relevant contributions, within the context of this paper. Second, the list of keywords was kept to a representative minimum by using the strings: “social business intelligence”, “social media AND business intelligence”. The keywords were applied to the title, abstract and keywords sections of scrutinized publications included in the SCOPUS database. The search includes publications in scientific journals, peer-reviewed conference proceedings and book chapters. The research focus of our approach led us to the decision to eliminate books and editorial reviews. We decided not to exclude publications in peer-reviewed conference proceedings, since SBI is a rather new and emerging scientific area and will be populated by more than a few first stage publications in the conference dissemination channel. Other types of publications such as notes and short surveys, are also excluded from the study.

The keyword search described above returned 131 papers published from 2006 to 2016, as shown in Table 1 and Table 2 below. These initial results show that contributions using SBI as a term are scarce (14) implying that, indeed, SBI is a scientific area in its infancy.

The collected papers were studied thoroughly with the objective to eliminate duplicates and then critically exclude the ones dealing with SBI issue superficially, fragmentally or not at all, in the case of the publications included in Table 3. Contributions that were included in the initial sample fulfilling the keyword string criteria but not directly dealing with the study subject were excluded from the database. Finally, a sum of 83 relevant papers was selected for in-depth, full-text examination with the objective to provide the reader with an overview of the main themes and trends covered by the relevant SBI literature.

Table 1 Search results for keyword string ‘social business intelligence’.

	Year of Publication										
Source Type	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**
Journal Paper										1	
Conference Paper							2	1	1	4	2
Book Chapter									1	2	
Sum = 14 papers							**2**	**1**	**2**	**7**	**2**

Table 2 Search results for keyword string “social media AND business intelligence”.

	Year of Publication										
Source Type	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**
Journal Paper				1	1	2	5	6	8	12	16
Conference Paper				1	2	5	8	8	11	13	11
Book Chapter							2	3	1		1
Sum = 117 papers				**2**	**3**	**7**	**15**	**17**	**20**	**25**	**28**

4. LITERATURE REVIEW

4.1 Descriptive Analysis

As stated in the previous section, the main body of literature identified comprises 83 papers. While 2006 is the first year of publication where contributions were sought, the first published papers found were from the year 2010, further validating the decision not to extend the study period prior to 2006. The allocation of the publications within the researched period (2006-2016) is presented in Figure 1.

The allocation of papers in the three source types, i.e. journal papers, papers in conference proceedings and book chapters, is presented in Figure 2. It is noted that contrary to what was expected there seems to be an even distribution between journal papers (44.6%) and conference

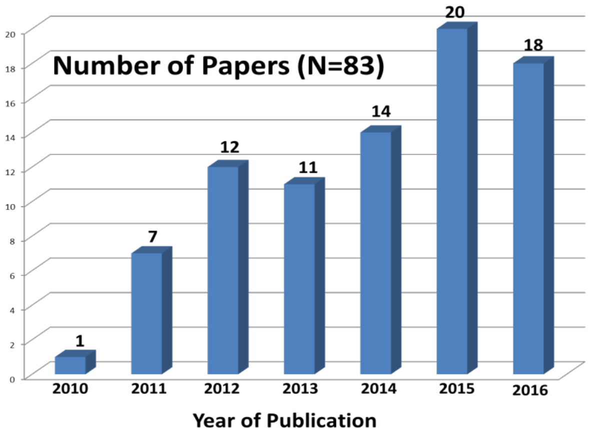


Figure 2 Distribution of publications per year across the study period.

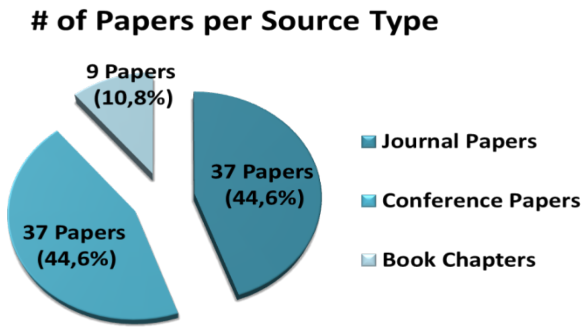


Figure 1 Allocation of publications in source types.

proceedings (44.6%), with book chapters corresponding to the smallest share (10.8%) of the work published on SBI between 2006 and 2016.

A closer look in the data set shows that a significant part of the publications (14 documents or 16.9%) originate from the Institute of Electrical and Electronics Engineer (IEEE) association (10 conference papers, 3 Journal papers and 1 book chapter). This is no surprise, since IEEE is a leading organization with a wide scientific area coverage including the technical background and information systems infrastructure necessary for SBI to be facilitated in companies and other organizations. In the same direction, Association for Computing Machinery (ACM) leads the item count when it comes to publications in peer reviewed scientific journals, with its cross-discipline journal entitled "ACM Transactions on Management Information Systems" leading the relative list with four publications. In Table 3, the number of papers per journal in the dataset is presented. The journals are presented in four categories depending on their main focus and thematic interest, i.e. information systems in management, information and computer science, social networks and miscellaneous. It is interesting to note the absence of any special issues dedicated to SBI and the scarcity of papers' appearance in more specific domains, such as social networks and journals.

4.2 Thematic Analysis

SBI, as evidenced from the descriptive statistics in the previous section, is a relatively new area, with the first publications referring explicitly to the term dating back to 2010. For those critically standing before the rapid emergence of the subject, SBI is nothing but the next logical step of BI evolution, providing enhanced collaborative capability in the decision-making process of an organization by adding the analytical capability pertaining to social media. For others, SBI is a BI paradigm revolution, especially when combined with the emergence of Big Data and the ever increasing variety, volume and velocity with which they arrive in front of business systems' queues for further processing in order to effectively support decision making. Needless to say that this duality of perspectives, coupled with the initial triggering of the term from the IT business area, has led to a plurality of terms describing SBI, still serving different business needs or marketing of IT products, thus creating confusion and reduced clarity on its definition. On the academic front, Dinter & Lorenz (2012), who according to our knowledge provide the single academic reference attempting to develop a framework of research in the SBI area, along the same lines as this paper, argue that lack of definition clarity for SBI might lie in the ongoing use of diverse related terms, coming mostly from industry literature freely accessible on the web. Zeng et al. (2010) provide one of the few available definitions of the term as a set of tools and techniques that "*derive actionable information from social media in context rich application settings, in order to develop corresponding decision-making frameworks and provide architectural designs and solution frameworks for existing and new business applications*". In this paper, the term SBI is explored in literature and used as the term of focus for the following study.

In that direction, thematic analysis of available literature in the SBI research field is organized into the four following distinct sections: the use of social media data in BI systems, SBI tools and techniques, BI applications in prevalent social media and finally, industry-specific SBI applications.

Table 3 Number of papers per scientific peer-reviewed journal.

Journal Title	# of Publications	Main Thematic Focus
ACM Transactions on Management Information Systems	**4**	Information Systems in Management
Decision Support Systems	**1**	
Journal of Enterprise Information Management	**1**	
Journal of the University of Pardubice	**1**	
Management Decision	**1**	
Production and Operations Management	**1**	
Technology Analysis & Strategic Management	**1**	
Journal of Decision Systems	**1**	
Journal of Destination Marketing & Management	**1**	
International Journal of Services Technology and Management	**1**	
Intelligent Systems in Accounting, Finance and Management	**1**	
Procedia Manufacturing	**1**	
	15	
Information Systems	**1**	Information & Computer Science
International Journal of Computer Technology and Applications	**1**	
Information Systems Frontiers	**1**	
Frontiers in Artificial Intelligence and Applications	**1**	
IEEE Computer Graphics and Applications	**1**	
IEEE Transactions on Knowledge and Data Engineering	**1**	
International Journal of Engineering & Technology	**1**	
IEEE Transactions on Visualization and Computer Graphics	**1**	
KSII Transactions on Internet and Information Systems	**1**	
Knowledge and Information Systems	**1**	
Scandinavian Journal of Information Systems	**1**	
Knowledge-Based Systems	**1**	
Mobile Networks and Applications	**1**	
Sensors	**1**	
Information Visualization	**1**	
Journal of Computer Information Systems	**1**	
Procedia Computer Science	**1**	
	17	
Journal of Internet Social Networking & Virtual Communities	**1**	Social Networks
International Journal of Sociotechnology and Knowledge Development	**1**	
Social Network Analysis and Mining	**1**	
	3	
The Decision Sciences Journal of Innovative Education	**1**	Miscellaneous
The Scientific World Journal	**1**	
	2	

4.2.1 The use of social media data in BI systems

Meredith and O'Donnell (2010; 2011) and Sathyanarayana et al. (2012) were among the first to detect the value of social networks in BI systems, beyond sales and marketing applications. They developed a framework to classify the social media functions that foster the Web 2.0 core concepts of user collaboration and contribution, and used it in order to exploit how it can "*create more effective and 'active' BI applications*". Shroff et al. (2011) introduced the term "*Enterprise Information Fusion*" to describe an emerging BI need across multiple industries, such as manufacturing, insurance and retail. The term includes the publicly available data, derived from social media, that can potentially be of immense business value for the enterprise ecosystem. On the same account, Ruhi (2014) attempted to outline the undeniable value of social media analytics, incorporated in a BI perspective. As he explains, the advantage of social media analytics in the business environment is twofold, as it can help organisations "*formulate and implement measurement techniques for deriving insights from social media interactions*" and, alongside "*evaluate the success of their own social media initiatives*". Wongthongtham and Abu-Salin (2015) emphasise the need of evaluation of traditional BI warehouses, which are more focused on handling structured internal enterprise data, in order to support the tremendous volume of valuable, yet unstructured, social media information, such as customer reviews and brand-related posts. Finally, Ram et al. (2016) conducted a survey in IT consultants and managers in various industry sectors of China, in order to prove the paradigm shift that social media have dictated in business strategies globally. With a semi-structured questionnaire, they managed to identify the critical issues in creating value through Big Data and social media analytics for BI systems.

Alongside the prevalence of social media in BI applications, several traditional business terms were redefined in academic literature in order to incorporate the new social trends, i.e. social customer relationship management (SCRM), digital marketing (Luo et al., 2015), voice of customer (VoC) and voice of market (VoM). Bachmann and Kantorova (2016) separate the original concept of CRM "*based rather on face-to-face and offline communication in the physical environment*", from social CRM which is mainly conducted "*through social networks and relationships within online communities*". Beverungen et al. (2014) argue that the global penetration of social networks constitutes a fertile ground for novel CRM strategies (Rosemann et al., 2012). After introducing the social CRM emerging field of research, the authors propose a framework to exploit Facebook data in CRM strategies and testify its applicability by building management reports for the retail industry. Berlanga et al. (2014, 2016) use advances in opinion mining techniques and sentiment analysis to describe the new opportunities arising in the VoM and VoC concepts in BI applications. As they explain, organisations can take advantage of the wealth of sentiment data in massive social media (e.g., social networks, product review blogs, forums) to 'listen' to their customers' needs and extract valuable business insights. In the same context, Lotfy et al. (2016) propose a framework to integrate customer opinion streams extracted from social media, with pre-existing corporate data, so as to constitute an integrated data warehouse. According to them, such a multidimensional data base "*can perform advanced analytical tasks and lead to better insights that would not have been possible to gain without this integration*".

Chan et al. (2015) deliberately examine the challenges faced by contemporary BI systems, associated with user-generated content (UGC) derived from social media communication channels. According to them, the available social data is not fully exploited for three main reasons: its unstructured format, its subjective nature and tremendous volume. On that account, they propose a systematic approach to social media data analysis, which counterbalances the aforementioned challenges and captures the real value of online social content for BI applications. Likewise, Tayouri (2015) also draws attention to risks associated with social media in the corporate environment, highlighting cyber security issues, such as fraud through social media activities, leakage of sensitive business information and damages to a firm's reputation. Hence, he suggests a consistent cyber security training framework supported by social media site monitoring tools, able to assist companies in building a robust SBI strategy by keeping track of correlated malicious activities and threats.

4.2.2 SBI Tools and Techniques

SBI tools and techniques is a predominant research field in academic literature. Data visualization tools; online analytical processing (OLAP); UGC and natural language processing (NLP) techniques; sentiment and opinion mining in the social media context; and user profiling and personalized marketing tools are some of the core thematic areas associated with contemporary SBI practices.

4.3 Visual Analytics

Zimmerman et al. (2015), Zimmerman & Vatrapu (2015) and Sigman et al., (2016) highlight the importance of visual analytics (VA) toolkits in assisting the interpretation of the unstructured data derived from social media into meaningful business or educational insights. Their research project provides a series of visual dashboards able to comprehensively project the analytics related to a brand and its marketing campaigns outcomes. The technical architecture and the specific characteristics of this set of dashboards is further explained and defined in '*Social Newsroom*', a prototype VA tool for SBI, which was developed to "*provide practitioners with user interfaces that can assist them in interpreting social media data and taking decisive actions*" (Zimmerman et al., 2015). Lu et al. (2014) proposed a VA toolkit able to handle "*noisy, unstructured data and use it for trend analysis and prediction*" in sales forecasting and advertisement analysis. Their data visualization tool was successfully applied in Twitter users, in order to predict movie revenue and ratings. Moreover, Pu et al. (2016) focused on the valuable geo-location data available in social media applications by introducing the '*Social Check-in Fingerprinting (Sci-Fin)*' tool which offers organisations "*the opportunity to capture and analyze users' spatial and temporal behaviors*" through social network check-in data. Respectively, Wen et al. (2016) suggested an alternative VA system, called '*SocialRadius'* that can interactively explore spatio-temporal features and check-in activities, in a variety of applications, ranging from BI applications to transportation and information recommendation systems. Meanwhile, Kucher et al. (2014; 2016) presented a VA tool for social media textual data that "*can be used to investigate stance phenomena and to refine the so-called stance markers collection*" with respect to sentiment and certainty. Lastly, Wu et al. (2014) introduced '*OpinionFlow'* a VA system detecting opinion propagation patterns and providing gleaned insights in government and BI applications.

4.4 OLAP techniques, UGC and NLP tools

Gallinucci et al. (2013; 2015) defined SBI as the "*discipline of combining corporate data with user-generated content (UGC) to let decision-makers improve their business, based on the trends perceived from the environment*". In order to enable contextual topics extraction and aggregation at different levels, they introduced 'Meta-Stars', a model based on UGC and real-time OLAP techniques. Golfarelli (2014) demonstrated an empirical application of a prototype demo of the model in real-world marketing campaigns, in order to prove its technical robustness and methodology. He furthermore presented the available OLAP solutions, for UGC analysis, that enable decision-makers analyze their business environment based on trends perceived from social media. Lin and Goh (2011) proposed a least-square (LS) algorithm to model sales performance and business value derived from social network data, by emphasizing the role of social marketer-generated content in influencing UGC sentiment and attitude. The authors actually suggest that there is a positive relationship between "*the richness of information embedded in both user-and marketer-generated content and firm sales performance*". Finally, Ferrara et al. (2014) provide a classification of the available UGC extraction tools in two main categories, namely the enterprise and social web data extraction (WDE) tools, through a structured literature review. In a natural language processing (NLP) context, Dey et al. (2011) discuss a series of methodologies that can be followed in order to "*obtain competitive intelligence from different types of web resources, including social media, using a wide array of text mining techniques*". As they explain, social media do not only provide valuable competition insights but also constitute an open forum where customers express their opinions about different brands' offerings. Sleem-Amer et al. (2012) introduce '*DoXa'*, a semantic search engine for the French language, with NLP capabilities and social BI application. Centering their work on two separate business cases, the authors explain how 'DoXa' can be applied to discover "*hidden patterns in social media data, using rich linguistic resources*". Lastly, Bjurstrom and Plachkinova (2015) propose a controlled natural language that does not require advanced technical skills and can be directly compiled into executable code, for automated social media data extraction.

4.5 Sentiment Analysis and Opinion Mining

Sentiment Analysis and Opinion Mining techniques are two research areas that attracted the academic interest from the early stages of introduction of social media as a powerful leverage for BI systems. Upon the rise of Web 2.0 and the increasing popularity of social network sites (SNS), Castellanos et al.

(2011), in collaboration with HP and its BI software solutions, introduced '*LCI*', a prototype sentiment analysis platform able to extract sentiment from textual data in real-time. The platform's interface consisted of multiple chart and visualization options that dynamically changed as soon as new data was ingested, exploiting state-of-the art sentiment analysis algorithms. A year later, Yang and Shih (2012) proposed a rule-based sentiment analysis (R-SA) technique "*to automatically discover interesting and effective rules capable of extracting product features or opinion sentences for a specific product feature*". That way, they offered a means of effective and real-time analysis of the tremendous volume of data, hidden in social media applications, regarding customer reviews about business offerings. In the same direction, Liu and Yang (2012) developed a buyer behavior prediction technique, using dynamic social network analysis and behavior pattern mining algorithms on e-commerce purchases and viral marketing applications. Qazi et al. (2014) focused on the suggestive type of customer reviews, found on online review forums, by combining machine learning techniques and sentiment analysis. Their findings suggested that sentiment analysis "*achieves maximum performance when employing additional preprocessing in the form of negation handling and target masking, combined with sentiment lexicons*". Later, Colombo et al. (2015) compared two novel methodologies for sentiment analysis with cross-industry application, by using secondary unstructured textual data from Twitter, Yelp and Cars.com, while Kim and Jeong (2015) applied their opinion mining methodology in online reviews about the oldest instant noodle snack in Korea. Finally, Nithya and Maheswari (2016) implemented a scoring system technique to identify the most promising features of a product offering, consisted of two rating attributes, namely the '*sentiment score'* and the '*feature score'*. Their technique provides managers with valuable insights regarding future demand, brand promotion and product penetration.

4.6 User Profiling and Personalized marketing tools

SBI tools and techniques for customer-centric marketing applications constitute another popular research field in academic circles. Personalized advertising messages, based on intelligent user profiling, is top priority for the contemporary business world, striving to survive in a highly competitive and globalized environment. Ranjan et al. (2014), using an association rules mining (ARM) algorithm, exploit social media data to locate tie-strength networks and active friends, in order to be used as a basis for targeted and relevant advertising campaigns. Yang and Chen (2014) introduce a novel profile expansion mechanism which enhances the effectiveness of personalized recommendation systems in social bookmarking sites to assist companies in developing "*effective service offerings that are better tailored to their customers' needs*" (Gronroos, 2008). In a more targeted approach about accurate profiling of social media users, Liu et al. (2014; 2015) develop '*HYDRA'*, a solution framework to identify linkages across multiple social networks and discover correlations between different user profiles. The authors argue that 'HYDRA' can be a profitable addition to existing BI solutions, as it was successfully implemented in a ten-million data base and correctly identified real user linkage, across seven dominant social network platforms, outperforming "*existing state-of-the art algorithms by at least 20% under different settings*". Finally, Yang and Chang (2015) highlight the knowledge gained from social tagging system (also known as folksonomy) as an invaluable asset for enhancement and upgrade of existing BI applications. On that account they employ Delicious, an established social bookmarking service, to construct "*a statistical-based thesaurus, which is then applied to support personalized document clustering*". Their empirical study indicated that such services improve the overall quality of SBI systems, and promote their efficiency in handling targeted marketing applications.

4.6.1 BI applications in prevalent social media

A fairly important percentage of existing academic literature on BI focuses on specific social media use-cases and their potential applicability on corresponding systems. Tools and techniques able to extract value added data from popular social network platforms, such as Twitter and Facebook blogs or websites containing customer review content, are among the most preferable research subjects.

4.7 Twitter

Rui and Whinston (2011) argue that Twitter hides a huge business potential as a base platform for BI applications, given its valuable structural information and the tremendous volume of data flows that are produced by its users in real-time. Within this context, they introduced a Twitter-based BI system for revenue forecasting, which was successfully implemented in movie box office revenue prediction, achieving remarkable results. Seebach et al. (2012) focused on the corporate reputation management area and how firms can use social media intelligence in order to handle reputation threats timely and effectively. By using sentiment and manual content analysis techniques on Twitter, regarding posts about a large American bank, they showed "*how social media might impact corporate reputation and what organizations can do to prepare themselves*". Lee et al. (2013) used Twitter as a real-time event detection system for crisis management and BI applications. Their proposed framework is able to detect emerging events from social network streams and "*accurately extract ontology entities associated with specific events for decision supporting applications*". O'Leary (2015) highlights 'Twitter Mining' as an invaluable asset for the majority of Fortune 100 companies. In his paper he reviews some of the most prevailing BI applications of Twitter data extraction on a prediction, discovery and informative basis. In the same year, Arora et al. (2015) applied sentiment analysis tools in order to investigate whether tweets provide a sufficient ground to gain useful insights on competitive brands, using the smart-phone industry. Their results showed that although Twitter data is rich regarding costumer sentiments, the exposure of different brands varies significantly making their comparison a rather ambiguous task. In a similar approach, Chilhare et al (2016) designed a marketing-driven competition analysis tool to "*recognize specific areas in which businesses are leading and lagging*". In their paper, they propose a methodology combining NLP techniques and sentiment benchmarks, in order to analyze and structure multilingual Twitter data for competitive FMCG companies into meaningful business insights. Sijtsma et al. (2016) introduced '*Tweet-viz*', an interactive tool to assist companies in actionable information extraction from unstructured textual Twitter data. In their paper, they prove that Twitter can provide BI systems, customer preferences, demographics and location data. Completing the Twitter BI academic cycle, Piccialli and Jung (2016) summarize the business-generated tweet content in three categories; namely informative, advertising or a hybrid of the two. According to their estimations, the hybrid approach increases customer engagement and promotes UGC activity with brand related content.

4.8 Facebook

According to scholars, Facebook, apart from being the most dominant social network on a global basis, contains such a high volume of UGC data that it could also turn to an alternative customer relationship management (CRM) platform, replacing traditional in-house corporate software. Bygstad and Presthus (2013), conducting a case study on two Scandinavian airliners' pages on Facebook during the ash crisis in April 2010, showed that CRM through the platform proved more effective in terms of dynamic interaction and customer engagement. Milolidakis et al. (2014), aiming to provide a generic framework for social media data extraction and transformation into meaningful business insights, used Facebook fan pages of three Greek communication service providers as their case study. According to their findings, Facebook includes data capturing of all the standard BI indicators, and moreover provides additional user sentiment information through artifacts features, such as the "like" button, that can turn to intelligent business statistics through visual excavation tools.

4.9 Blogs and Micro-Blogs

Banerjee and Agarwal (2012) used a nature-inspired theory to model collective user behavior from blog-originated data in order to explore its application on BI systems. Based on swarm intelligence, "*where the goal is to accurately model and predict the future behavior of a large population after observing their interactions during a training phase*", they concluded in promising results about blog value in trend prediction applications. Meanwhile, Kulkarni et al. (2013) draw attention on the importance of social media brand propagation enablers. On that basis, they study the degree of customer engagement through blog contents and the corresponding analytics for BI systems. Obradović et al. (2013), in the context of the '*Social Media Miner*' project combined textual analysis methods with a blog processing technique to "*aggregate blog articles of a specific domain from multiple search services, analyze the social* *authorities of articles and blogs and monitor the attention they receive over time*", in order to provide a highly automated BI tool. Lastly, Jingjing et al. (2013), using as a reference the Chinese micro-blogging platform '*Sina Weibo*', conduct a social influence analysis to discover "*information retrieval, recommendations and businesses intelligence opportunities*". According to their findings, their proposed framework can overcome difficulties related to volume and complexity found on micro-blogging platforms and can find numerous applications in BI systems.

4.10 Amazon.com

Social media, based on principles and technologies deriving from the user-centered Web 2.0, constitute by definition an open platform where users can express their sentiments, share their knowledge and build a social environment. Within this context, consumers exchange their opinions towards different brands, share their experiences through word-of-mouth (WOM) and provide their own reviews. The importance of social activity related to brand offerings and the added-value of publicly available customer reviews has naturally attracted the interest of the business and academic world. Zhang and Chen (2012) studied the business impact of social media and UGC in sales and marketing, by applying text mining techniques and a set of innovative metrics focusing on customer reviews on two popular e-commerce websites, namely BN.com (Barnes&Noble) and Amazon.com. According to their findings, user-generated reviews have serious effects on product sales and should be consistently processed by BI systems, through carefully selected measures. Similarly, Ngo-Ye and Sinha (2012) argue that customer-generated reviews in e-marketplaces "*are playing an increasingly important role in disseminating information, facilitating trust, and promoting commerce*". On that account, they developed an Amazon.com based tool to automatically identify the most important reviews and provide meaningful customer feedback. Finally, Zhang et al. (2013) in an attempt to further explain how WOM is affecting product sales, they combined network analysis with textual sentiment mining techniques to build product-comparison networks consisted of customer reviews. Their empirical study on Amazon.com suggests that it is imperative for firms to understand and manipulate the WOM process taking place in social media, in order to survive in the increasingly competitive online landscape.

4.10.1 Industry-specific SBI applications

The integration of social media analytics in BI systems is a need soon realized both by organisations and academia. SBI tools and techniques are not limited in a specific area but have rather a cross-industry application, a fact that is clearly reflected in the existing literature. Heijnen et al. (2013) argue that the potential of social medial data is invaluable for multiple facets of BI systems, yet it is "*largely unused by companies, and it remains unclear what data can be useful for which industry sectors*". Their findings indicate that key performance indicators typically differ between industry sectors and therefore SBI metrics should accordingly adapt to their corresponding needs. The need to approach the matter from industry-specific perspectives led to a series of academic publications focusing on distinct sectors: education, automotive, pharmaceutical, cosmetics, tourism, fashion, government and politics.

4.11 Education

Moedeen and Jeerooburkhan (2016) focus on the higher education sector to explore how "*social media strategies can be aligned with business strategies to help universities gain a competitive edge*". Using the Facebook page of a university as their case study, they argue that higher education organizations pay attention solely to advertising and reputation management aspects, while neglecting other business objectives that could be met through a holistic SBI application.

4.12 Government and Politics

Bendler et al. (2014) associate static environmental characteristics with dynamic user-generated content from social media to explain and predict criminal activity in metropolitan areas. By combining traditional statistics, such as zero-inflated Poisson regression and geographically weighted regression with social media data, they provide a framework that enhances the accuracy of criminal activity forecasting. Meanwhile, Chung et al. (2014) developed an approach to pinpoint opinion leaders in social networking sites that could be approached by policy makers to collaborate and "*bring about change in the communities and the general public welfare*". In a more generic approach, Golfarelli

(2014; 2015) studies SBI options in politics. According to him, processing of user-generated content through a robust SBI system could prove invaluable for political entities in order to align their governmental decisions with environmental trends and public opinion. Finally, Beigi et al. (2016) explore crisis and disaster management through sentiment analysis and social media visual analytics. According to them, individual posts in social media about natural disasters and emergencies can be used as inputs in governmental SBI systems "*to improve situational awareness and crisis management (…) while assisting in locating people who are in specific need during emergency situations*".

4.13 Automotive

Abrahams et al. (2012) introduce a decision support technique for vehicle quality management designed to identify and prioritise automotive defects, deriving from reviews in vehicle enthusiasts' online forums. They suggest that conventional sentiment analysis does not suffice to efficiently detect customer complaints and therefore, BI systems should incorporate advanced text mining algorithms specifically designed for social media applications. Baur et al. (2015) also focus on the vehicle industry by exploring Chinese auto forums as a new proactive means of market research. According to them, although the increasing popularity of social media offers a fertile ground for novel marketing techniques, there is a number of arising challenges to be confronted, namely the tremendous volume of posts, their unstructured format and the wide range of user languages requiring complex natural language processing techniques. On that basis, they propose '*MarketMiner'* a novel framework for "*search, integration, and analysis of cross-language user-generated content*", specifically designed for the competitive automotive sector. One year later, Baur (2016) examines alternative applications for 'MarketMiner' in public administrative bodies and commercial firms. His results indicate that the tool can significantly improve the processing of multi-source and multi-language social media generated content and apply to cross-industry SBI systems.

4.14 Pharmaceutical

According to Basset et al. (2012) the social media sphere is a challenging environment for the pharmaceutical industry, as it is associated with a number of ethical and legal issues imposed by governments globally. However, SBI systems can prove valuable to such an antagonistic sector mainly for marketing, customer relationship management and competition monitoring applications. Bell and Shirzad (2013a; 2013b) propose a social media data extraction model to assist pharmaceutical companies to effectively position themselves in new marketplaces. According to them, social media networks offer a channel of communication for business-to-business environments and can enable companies to connect with all the actors of their value chain (i.e. customers, partners and even competitors) on a real-time, global basis. Finally, He et al. (2016) using the three biggest drugstore chains in US as their case-study, suggest a model for competitive strategy formulation by applying quantitative analysis, sentiment analysis and text-mining techniques in social media UGC content. Their findings indicate that such tools can prove invaluable if adopted by existing BI systems.

4.15 Tourism

Online social networks, and Web 2.0 applications in general, are rapidly becoming a significant marketing channel for the tourist industry which is challenged by new and emerging business models utilizing social media and other crowd sourcing and shared economy applications, such as Airbnb. In this new and turbulent environment, social BI can be a source of critical competitive advantage in a very demanding and customer-service intensive industry such as tourism. In that direction, Palacios-Marqués et al. (2015) study the effect of online social networks on firm performance and explore ways of adding value to established market competences. The authors conduct a large survey in one of the world's largest tourist destination, Spain, with the participation of top managers from 197 four- and five-star hospitality firms. Their results show that social BI has a significant positive relationship with firm performance by enhancing market intelligence and knowledge management competences, thus leading to the acquisition of a significant advantage over the competition. Remaining in the same geographic territory but penetrating one layer deeper in the social BI area, Marine-Roig and Clave (2015) study the usefulness and applicability of big data analytics for the industry and specifically for a smart city tourist destination, Barcelona. The authors study the online image of the city by analyzing more than a hundred thousand travel blogs and online travel reviews by people who have visited the destination in the last decade. By extracting BI through these large volumes of user generated content, the authors provide an efficient decision support tool for industry executives and city officials to develop and evaluate competition, marketing, branding and positioning strategies and policies, which will enhance the city's image as a smart tourist destination.

4.16 Fashion and Luxury

The fashion and luxury products industry has for many years resisted the adoption of the e-commerce channel, since they associated anything digital with malpractices such as discounting, counterfeiting and brand dilution. This is not the case anymore and currently emblematic brands, such as Ferragamo, have entered the e-market arena, which according to a report from McKinsey and Altagamma (2015) has reached €14 billion in 2014, a 50% increase from 2013. This change has created the need for managing user-generated content in order to better understand customer profiles, identify preferences and determine trends, with the latter playing a crucial role for product development of companies in the fashion industry. In that direction, Petychakis et al. (2016) turn their research focus on a very important aspect of social media analysis, which is the identification of opinion makers within the social media ecosystem, the monitoring of their behavior and the extraction of targeted campaigns utilizing their media presence. The authors present a platform providing marketers and product designers with data analytic services for influencer identification and trend analysis and evaluate it in a single case from the fashion industry. Fourati-Jamoussi (2015) explores the concept of e-reputation by applying BI practices to analyze the social media presence of four companies from the organic cosmetics industry. The author attempts to compare the reputation of the participating brands by using different monitoring tools, conducts user profiling for each brand and finally proposes recommendations for enhancing marketing strategies.

5. CONCLUSIONS

In this paper 83 papers, which were published in the period from 2006 to 2016 dealing with SBI concept, management, tools and applications, were collected. The review of these papers and the analysis of their content, presented in the previous chapter, produced useful information, in order to synthesize a comprehensive research agenda for SBI including major directions and identified shortcomings that seem to shape the future of research in this area. The core focus of the research, as expected, seems to be the unearthing of the currently unused, to its full potential, value of SBI and put it to good use for the benefit of businesses and organizations around the globe. In that direction, academic literature in the novel research field of SBI is essentially developed around three main pillars of research orientation.

The first pillar attempts to provide answers to the question '*What is SBI and how can it help a business or organization*", putting SBI's business validation and real-life applications in the epicenter of research, thus given the title 'business descriptive'. Papers in this pillar are attempting to highlight the prevalent acceptance of social media as a source of business value and the parallel expanded usage of BI systems through social media data for multiple operations within companies, in a variety of industry-specific applications. In doing so, they provide mostly definitions, methods, models and frameworks, which support a wide range of corporate activities, spanning but not limited to strategic decision-making functions, business processes' optimization, operational efficiency improvement and revenue management. Within this pillar, one can identify two discrete waves of publications that can be organized together based on their focus and objectives.

The first identified wave of publications within the business descriptive pillar deals mostly with determining the current status quo of BI in contemporary organizations and provides means of expanding its reach through the exploitation of social media. The first step in this direction is the identification and validation of social media potential and functionalities to act as a consistent BI decision-making support tool through solid argumentation and empirical tools like surveying experts in interested business areas. At the same time, the second wave of publications attempts to deal with SBI by exploring the enhanced capabilities that it gives to traditional BI systems and how these can be rethought and restructured in order to be ready to absorb and process the abundance and large variety of data that social media produce. What is interesting at this point is the determination of SBI usefulness and transformative impact on other established business functions such as marketing and customer relationship management (CRM). The introduction of novel marketing and CRM strategies such as social CRM, VoM and VoC as a result of information harnessed by SBI practices is explored in depth by many publications and specific algorithmic SBI techniques and tools, e.g. opinion mining, sentiment mining, are mentioned as playing a critical role for business success and competitive advantage. Finally, the main barriers/shortcomings identified in this pillar of literature are the following:

- Probably the most important issue identified is that of data security and privacy. There are major concerns for all levels of data usage, i.e. data creators (users), data suppliers (e.g. Facebook or Telco companies) and businesses in need of the data. What makes the situation even more complicated is the fragmentation of legislation between continents and nations, which make compliance a cumbersome and sensitive task, especially in the case of companies operating at a global level.

- The second most important issue identified in this pillar of literature is data governance by businesses. In other words, the ability of companies to streamline their processes and systems in order to provide more accurate information, achieve increased visibility and in essence better analytics. There seems to be a consensus that much more is needed to be accomplished in this area.

- Finally, the third prominent issue identified in this pillar is process governance. The huge impact of social media data on current established business processes and its transformative effect on every-day operations, coupled with the need for the use of more advanced analytical systems, creates the need for research on business process management and reevaluation of traditional processes and their efficient transformation.

The second pillar attempts to reveal 'under the hood' knowledge and answer the question "*How does SBI work*". It sets technical implementation in the epicenter of research, thus this pillar is given the title, 'technical descriptive'. Papers in this pillar provide mostly technical information on algorithms, techniques and tools that are used in order for SBI to process social media data and produce meaningful information to be used directly or passed for further processing by traditional BI systems. The prevalent techniques, which seem to dominate the research interest in this pillar are those dealing with three major issues of SBI at the technical level: user profiling, user (customer) voice translation into actionable information and data visualization. The main shortcomings identified in this pillar of literature are the following:

- There is an increased demand for new AI algorithms for the automation of the user generated content extraction and translation procedure. Current algorithmic efforts in research are many. Still their validation in actual commercial environments does not commensurate with the materialized research. The need for a switch towards an AI based, data-scientist agnostic SBI process is evident in the literature.

- User profiling and the underlying targeted marketing and personalized recommender systems are very important issues in the SBI literature especially for companies that are forced to enter the paid advertising arena by increased competition and the need to sustain profitability. Although profiling models and algorithms present a rapid increase in numbers and variety of approaches there are still several unexplored areas in profiling that need intensified research and investments.

- Data visualization has seen many advances in the last few years with the emergence of the dashboard logic in data presentation and display. Although there is a fair number of social media tools already providing services like data collection, aggregation and analysis into key performance indicators, there is still a deficiency in visualizations, especially when it comes to standards and design principles, thus making the support from data scientists and supplementary systems mandatory.

Finally, the third trend attempts to answer the question "*Does SBI work in real life?*" Real-life cases of SBI applications in practice are the focus of research in this pillar, which thus is given the title "case descriptive". Two discrete waves of publications can be identified. The first focuses on industry-specific applications and describes how SBI can provide valuable services for businesses operating in these industries. In doing so, papers in this pillar explore successful applications of SBI in real business cases, highlighting the cross-industrial nature of SBI and its potential impact for a variety of industries and governmental organizations. Specifically, they provide focus on the impact of SBI in traditional business models and processes and its operational fit in order to support industry-particular requirements. The second wave of publications includes papers focusing on specific social media use-cases, with Twitter and Facebook being the platforms most widely used as data providers and application test beds. Tools and techniques able to extract value added data from popular social network platforms, blogs or websites containing customer review content, are among the most preferable research subjects. The main shortcomings identified in this pillar of literature are the following:

- Utilizing SBI to support real-life cases is a cumbersome task demanding a holistic approach, including technological and organizational aspects, leading to a complex transition requiring high executive competences supported by a global strategy. This is not the case dealt within the publications studied in this pillar. Empirical evidence provided is rather fragmented and cases seem isolated from the business 'big picture', while connection with 'bottom line' metrics is loose.

- There is a, to some point justifiable, strong focus of SBI research on social networking giants, like Facebook and Twitter. Still, there is an abundance of social networking sites and other emerging social media business models like Snapchat, Vine and Reddit for example, for which the possibility of more open data extraction and enhanced algorithmic testing could take place, that are currently not sufficiently explored.

At this point, a research agenda can be formed including eight discrete research directions, each one dealing with the shortcomings identified in literature and discussed previously in this section. In Table 4, a summary of the literature review's main findings is presented. The three main pillars' research offerings are shortly described and specific publications are assigned to each one of the pillars in accordance with their number in the reference list at the end of this assignment. The eight research directions comprising the future research agenda for SBI are categorized per pillar and presented in Table 4.

Finally, it has to be noted that adoption of this paper's findings should take into account the inherent limitations of this study, which are:

- The big difference between current and published capabilities of academia, especially coupled with the fast pace of the SBI scientific field. The author is certain that several research efforts providing innovative approaches and empirical use

Table 4 SBI future research directions.

	TITLE (MAIN RESEARCH OFFERINGS)	**RESEARCH DIRECTIONS**		
FIRST PILLAR	**Business Descriptive** (Definitions, Methods, Models & Frameworks)	**RD1:** Data Security & Privacy	**RD2:** Data Governance	**RD3:** Process Governance
SECOND PILLAR	**Technical Descriptive** (Algorithms, Techniques and Tools)	**RD4:** Improvement / Development of new AI Algorithms - SBI Process Automation	**RD5:** User Profiling	**RD6:** Data Visualization
THIRD PILLAR	**Case Descriptive** (Empirical Evidence / Industry Focus & Social media Focus)	**RD7:** Holistic SBI Approaches - Enhanced Validation	**RD8:** Extend Research Coverage in Social Media	

cases highlighting novel applications of techniques do exist, that are either in development or already finished but yet unpublished. Unfortunately, the academic publishing pipeline has a lead time of six to eighteen months in some cases and work in progress papers are relatively low in numbers, thus creating problems to researchers who conduct a literature review.

- Furthermore, significant research on SBI is done by or on behalf of big players in this area, such as social media platforms, big advertising companies and global brands. These studies are based on home-grown methodologies, use proprietary tools and perhaps focal datasets and thus never made public, making the task of the researcher who conducts the review even more difficult.

- Finally, the reader, before adopting the results of this study, has to consider its methodological limitations, related to the selection of the academic library, i.e. Elsevier's SCOPUS, the inclusion of specific source types, i.e. peer reviewed journal papers, conference proceedings and book chapters and finally the selection of the search keywords for conducting the review.

6. REFERENCES

Abrahams, A.S., Jiao, J., Wang, G.A. and Fan, W. 2012. Vehicle defect discovery from social media. *Decision Support Systems*, 54(1), 87-97.

Arora, D., Li, K.F. and Neville, S.W. 2015. Consumers' sentiment analysis of popular phone brands and operating system preference using Twitter data: A feasibility study. In: *Proceedings of Advanced Information Networking and Applications (AINA) IEEE 29th International Conference*, pp. 680-686.

Bachmann, P. and Kantorová, K. 2016. From customer orientation to social CRM. New insights from Central Europe. *Scientific papers of the University of Pardubice*, Series D, Faculty of Economics and Administration, 36/2016.

Banerjee, S. and Agarwal, N. 2012. Analyzing collective behavior from blogs using swarm intelligence. *Knowledge and Information Systems*, 33(3), 523-547.

Basset, H., Stuart, D. and Silbe, D. 2012. *From Science 2.0 to Pharma 3.0 Semantic Search and Social Media in the Pharmaceutical Industry and Stm Publishing*. A volume in Chandos Publishing Social Media Series.

Baur, A., Lipenkova, J., Bühler, J. and Bick, M. 2015. *A Novel Design Science Approach for Integrating Chinese User-Generated Content in Non-Chinese Market Intelligence.*

Baur, A.W. (016. Harnessing the social web to enhance insights into people's opinions in business, government and public administration. *Information Systems Frontiers*, pp.1-21.

Beigi, G., Hu, X., Maciejewski, R. and Liu, H. 2016. An overview of sentiment analysis in social media and its applications in disaster relief. *Sentiment Analysis and Ontology Engineering*, pp. 313-340, Springer International Publishing.

Bell, D. and Shirzad, S. R. 2013. Social media business intelligence: A pharmaceutical domain analysis study. *International Journal of Sociotechnology and Knowledge Development (IJSKD)*, 5(3), pp. 51-73.

Bell, D. and Shirzad, S.R. 2013. Social Media Domain Analysis (SoMeDoA)-A Pharmaceutical Study. *WEBIST*, pp. 561-570.

Bendler, J., Ratku, A. and Neumann, D. 2014. Crime mapping through geo-spatial social media activity. In: *Proceedings of 35th International Conference on Information Systems*, Auckland 2014.

Berlanga, R., Aramburu, M.J., Llidó, D.M. and García-Moya, L. 2014. Towards a semantic data infrastructure for social business intelligence. *New Trends in Databases and Information Systems*, pp. 319-327, Springer International Publishing.

Berlanga, R., García-Moya, L., Nebot, V., Aramburu, M.J., Sanz, I. and Llidó, D.M. 2016. Slod-bi: An open data infrastructure for enabling social business intelligence. *Big Data: Concepts, Methodologies, Tools, and Applications*, pp. 1784-1813, IGI Global.

Beverungen, D., Eggert, M., Voigt, M. and Rosemann, M. 2014. Augmenting Analytical CRM Strategies with Social BI. *Digital Arts and Entertainment: Concepts, Methodologies,* *Tools, and Applications*, pp. 558-576, IGI Global.

Bjurstrom, S. and Plachkinova, M. 2015. *Sentiment Analysis Methodology for Social Web Intelligence*.

Bygstad, B. and Presthus, W. 2013. Social Media as CRM? How two airline companies used Facebook during the “Ash Crisis” in 2010. *Scandinavian Journal of Information Systems,* 25(1), 3.

Castellanos, M., Dayal, U., Hsu, M., Ghosh, R., Dekhil, M., Lu, Y., ... & Schreiman, M. 2011. LCI: a social channel analysis platform for live customer intelligence. In Proceedings of the 2011 ACM SIGMOD International Conference on Management of data (pp. 1049-1058). ACM.

Chan, H.K., Wang, X., Lacka, E. and Zhang, M. 2015. A Mixed-Method Approach to Extracting the Value of Social Media Data. *Production and Operations Management*.

Chaudhuri, S., Dayal, U. and Narasayya, V. 2011. An overview of business intelligence technology. *Communications of the ACM*, 54(8), 88-98.

Chen, H., Chiang, R.H. and Storey, V.C. 2012. Business intelligence and analytics: From big data to big impact. *MIS quarterly*, 36(4), 1165-1188. ISO 690

Chilhare, Y.R., Londhe, D.D. and Competiti, E.M. 2016. Competitive Analytics Framework on Bilingual Da Bilingual Dataset of Amazon Food Product. *IJCTA*, 9(21), pp. 179-189.

Chung, W., Zeng, D. and O'Hanlon, N. 2014. Identifying influential users in social media: A study of US immigration reform. In: *Proceedings of the 20th Americas Conference on Information Systems*, Savannah, 2014.

Colombo, C., Grech, J.P. and Pace, G.J. 2015. *A controlled natural language for business intelligence monitoring*. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 9103, pp. 300-306.

Dey, L., Haque, S.M., Khurdiya, A. and Shroff, G. 2011. Acquiring competitive intelligence from social media. In: *Proceedings of the 2011 joint workshop on multilingual OCR and analytics for noisy unstructured text data*, p. 3. ACM.

Diamantopoulou, V., Charalabidis, Y., Loukis, E., Triantafillou, A., Sebou, G. Foley, P., Deluca, A., Wiseman, I. and Koutzeris, T. 2010. *Categorization of Web 2.0 Social Media and Stakeholder Characteristics*. Nomad Project. EU. pp.19. Available at: http://www.padgets.eu/Downloads/Deliverables/tabid/75/ctl/Versions/mid/623/Itemid/56/Default.aspx [Accessed 2 March 2017]

Dinter, B. and Lorenz, A. 2012. Social business intelligence: a literature review and research agenda. In: *Proceedings of the 33rd International Conference on Information Systems*, Orlando 2012.

Fan, S., Lau, R.Y. and Zhao, J.L. 2015. Demystifying big data analytics for business intelligence through the lens of marketing mix. *Big Data Research*, 2(1), 28-32.

Ferrara, E., De Meo, P., Fiumara, G. and Baumgartner, R. 2014. Web data extraction, applications and techniques: A survey. *Knowledge-Based Systems*, 70, 301-323.

Fourati-Jamoussi, F. 2015. E-reputation: A case study of organic cosmetics in social media. In: *Proceedings of the Information Systems and Economic Intelligence (SIIE) 6th International Conference*, pp. 125-132, IEEE.

Gallinucci, E., Golfarelli, M. and Rizzi, S. 2013. Meta-stars: multidimensional modeling for social business intelligence. In: *Proceedings of the 16th international workshop on Data warehousing and OLAP*, pp. 11-18, ACM.

Gallinucci, E., Golfarelli, M., & Rizzi, S. 2015. Advanced topic modeling for social business intelligence. *Information Systems*, 53, 87-106.

Golfarelli, M. 2014. Social business intelligence: OLAP applied to user generated contents. In: *Proceedings of the e-Business (ICE-B) 11th International Conference*, pp. IS-11, IEEE.

Golfarelli, M. 2015. Design Issues in Social Business Intelligence Projects. In European Business Intelligence Summer School (pp. 62-86). Springer International Publishing.

Gronroos, C. 2008. Service logic revisited: Who creates value? And who co-creates? *European Business Review*, Vol. 20, No. 4, pp. 298–314.

Hart C. 1998. *Doing a Literature Review*. Sage Publications, London

He, W., Tian, X., Chen, Y. and Chong, D. 2016. Actionable social media competitive analytics for understanding customer experiences. *Journal of Computer Information Systems*, 56(2), 145-155.

Heijnen, J., De Reuver, M., Bouwman, H., Warnier, M. and Horlings, H. 2013. Social media data relevant for measuring key performance indicators? A content analysis approach. In: *Proceedings of the International Conference on Electronic Commerce*, pp. 74-84, Springer Berlin Heidelberg.

Jingjing, W., Changhong, T., Xiangwen, L. and Guolong, C. 2013. Mining Social Influence in Microblogging via Tensor Factorization Approach. In: *Proceedings of Cloud Computing and Big Data (CloudCom-Asia),* December 2013 International Conference, pp. 583-591, IEEE.

Kaplan, A. M. and Haenlein, M. 2010. Users of the world, unite! The challenges and opportunities of Social Media. *Business horizons*, 53(1), 59-68.

Keele, S. 2007. Guidelines for performing systematic literature reviews in software engineering. *In Technical report*, Ver. 2.3 EBSE Technical Report. EBSE.

Kim, Y. and Jeong, S. R. 2015. Opinion-Mining Methodology for Social Media Analytics. *TIIS*, 9(1), 391-406.

Kucher, K., Kerren, A., Paradis, C. and Sahlgren, M. 2014. Visual analysis of stance markers in online social media. In: *Proceedings of Visual Analytics Science and Technology (VAST)*, 2014 IEEE Conference, pp. 259-260, IEEE.

Kucher, K., Schamp-Bjerede, T., Kerren, A., Paradis, C. and Sahlgren, M. 2016. Visual analysis of online social media to open up the investigation of stance phenomena. *Information Visualization*, 15(2), 93-116.

Kulkarni, A. V., Joseph, S., Raman, R., Bharathi, V., Goswami, A. and Kelkar, B. 2013. Blog Content and User Engagement-An Insight Using Statistical Analysis. *International Journal of Engineering and Technology*, 5(3), pp. 2719-2733.

Lee, C., Wu, C., Wen, W. and Yang, H. 2013. Construction of an event ontology model using a stream mining approach on social media. *In: Proceedings of the 28th International Conference on Computers and Their Applications*, 2013, CATA 2013, pp.249-254.

Lin, Z. and Goh, K. Y. 2011. Measuring the business value of online social media content for marketers. In: *Proceedings of the 32nd International Conference on Information Systems*, Shanghai.

Liu, S., Wang, S. and Zhu, F. 2015. Structured learning from heterogeneous behavior for social identity linkage. *IEEE Transactions on Knowledge and Data Engineering*, 27(7), 2005-2019.

Liu, S., Wang, S., Zhu, F., Zhang, J. and Krishnan, R. 2014. Hydra: Large-scale social identity linkage via heterogeneous behavior modeling. In: *Proceedings of the 2014 ACM SIGMOD international conference on Management of data*, pp. 51-62, ACM.

Liu, X. and Yang, J. 2012. Social buying met network modeling and analysis. *International Journal of Services Technology and Management*, 18 (1- 2), 46-60.

Lotfy, A., El Tazi, N and El Gamal, N. 2016. SCI-F: Social-Corporate Data Integration Framework. In: *Proceedings of the 20th International Database Engineering & Applications Symposium*, June 2016, pp. 328-333, ACM.

Lu, Y., Wang, F. and Maciejewski, R. 2014. Business intelligence from social media: A study from the vast box office challenge. *IEEE computer graphics and applications*, 34(5), 58-69.

Luhn, H. P. 1958. A business intelligence system. *IBM Journal of Research and Development*, 2,14-31

Luo, J., Pan, X. and Zhu, X. 2015. Identifying digital traces for business marketing through topic probabilistic model. *Technology Analysis & Strategic Management*, 27(10), 1176-1192.

Marine-Roig, E., & Clavé, S. A. 2015. Tourism analytics with massive user-generated content: A case study of Barcelona. *Journal of Destination Marketing & Management*, 4(3), 162-172.

McKinsey and Altagamma 2015. *Digital inside: Get wired for the ultimate luxury experience.* Available at: https://www.mckinsey.de/files/dle-2015-global-report.pdf [Accessed 5 March 2017]

Meredith, R. and O'Donnell, P. A. 2010. A Functional Model of Social Media and its Application to Business Intelligence. In: *Proceedings of the 2010 conference on Bridging the Socio-technical Gap in Decision Support Systems: Challenges for the Next Decade*, August 2010, pp. 129-140, IOS Press, Netherlands.

Meredith, R. and O'Donnell, P. A. 2011. A framework for understanding the role of social media in business intelligence systems. *Journal of Decision Systems*, 20(3), 263-282.

Milolidakis, G., Akoumianakis, D. and Kimble, C. 2014. Digital traces for business intelligence: A case study of mobile telecoms service brands in Greece. *Journal of Enterprise Information Management*, 27(1), 66-98.

Moedeen, B. W. and Jeerooburkhan, A.S. 2016. Evaluating the strategic role of Social Media Analytics to gain business intelligence in Higher Education Institutions. In: *Proceedings of Emerging Technologies and Innovative Business Practices for the Transformation of Societies (EmergiTech),* IEEE International Conference, pp. 303-308.

Ngo-Ye, T. L. and Sinha, A.P. 2012. Analyzing online review helpfulness using a regressional ReliefF-enhanced text mining method. *ACM Transactions on Management Information Systems (TMIS)*, 3(2), 10.

Nithya, R. and Maheswari, D. 2016. Correlation of feature score to overall sentiment score for identifying the promising features. In: *Proceedings of Computer Communication and Informatics (ICCCI) International Conference*, January 2016, pp. 1-5, IEEE.

O'Leary, D. E. 2015. Twitter Mining for Discovery, Prediction and Causality: Applications and Methodologies. *Intelligent Systems in Accounting, Finance and Management*, 22(3), 227-247.

Obradović, D., Baumann, S. and Dengel, A. 2013. A social network analysis and mining methodology for the monitoring of specific domains in the blogosphere. *Social Network Analysis and Mining*, 3(2), 221-232.

Olszak, C.M. 2016. Toward better understanding and use of Business Intelligence in organizations. *Information Systems Management*, 33(2), 105-123.

Palacios-Marqués, D., Merigó, J. M. and Soto-Acosta, P. 2015. Online social networks as an enabler of innovation in organizations. *Management Decision*, 53(9), 1906-1920.

Petychakis, M., Biliri, E., Arvanitakis, A., Michalitsi-Psarrou, A., Kokkinakos, P., Lampathaki, F. and Askounis, D. 2016. *Detecting Influencing Behaviour for Product-Service Design through Big Data Intelligence in Manufacturing. In: Proceedings of Working Conference on Virtual Enterprises*, pp. 361-369, Springer International Publishing.

Piccialli, F. and Jung, J. E. 2016. Understanding Customer Experience Diffusion on Social Networking Services by Big Data Analytics. *Mobile Networks and Applications*, 1-8.

Ponis, S. T., & Christou, I. T. 2013. Competitive intelligence for SMEs: a web-based decision support system. International Journal of Business Information Systems, 12(3), 243-258.

Pu, J., Teng, Z., Gong, R., Wen, C. and Xu, Y. 2016. Sci-Fin: Visual Mining Spatial and Temporal Behavior Features from Social Media. *Sensors*, 16(12), 2194.

Qazi, A., Raj, R.G., Tahir, M., Cambria, E. and Syed, K.B.S. 2014. Enhancing business intelligence by means of suggestive reviews. *The Scientific World Journal*, 2014.

Ram, J., Zhang, C. and Koronios, A. 2016. The Implications of Big Data Analytics on Business Intelligence: A Qualitative Study in China. *Procedia Computer Science*, 87, 221-226.

Ranjan, J. 2009. Business intelligence: Concepts, components, techniques and benefits. *Journal of Theoretical and Applied Information Technology*, 9(1), 60-70.

Ranjan, R., Vyas, D. and Guntoju, D. P. 2014. Balancing the trade-off between privacy and profitability in Social Media using NMSANT. In: *Proceedings of Advance Computing Conference (IACC)*, 2014 IEEE International, pp. 477-483, IEEE.

Rosemann, M., Eggert, M., Voigt, M. and Beverungen, D. 2012. Leveraging social network data for analytical CRM strategies: the introduction of social BI. In: *Proceedings of the 20th European Conference on Information Systems (ECIS) 2012,* AIS Electronic Library (AISeL).

Ruhi, U. 2014. Social Media Analytics as a business intelligence practice: current landscape & future prospects. *Journal of Internet Social Networking & Virtual Communities*, 2014.

Rui, H., & Whinston, A. 2011. Designing a social-broadcasting-based business intelligence system. *ACM Transactions on Management Information Systems (TMIS)*, 2(4), 22.

Sathyanarayana, P., Tran, P.N.K., Meredith, R. and O'Donnell, P. A. 2012. Towards a Protocol to Measure the Social Media Affordances of Web Sites and Business Intelligence Systems. *DSS*, pp. 317-322.

Seebach, C., Beck, R. and Denisova, O. 2012. Sensing Social Media for Corporate Reputation Management: a Business Agility Perspective. *ECIS*, p. 140.

Shroff, G., Agarwal, P. and Dey, L. 2011. Enterprise information fusion for real-time business intelligence. In: *Proceedings of the 14th International Conference*, Information Fusion (FUSION), pp. 1-8, IEEE.

Sigman, B. P., Garr, W., Pongsajapan, R., Selvanadin, M., McWilliams, M. and Bolling, K. 2016. Visualization of Twitter Data in the Classroom. *Decision Sciences Journal of Innovative Education*, 14(4), 362-381.

Sijtsma, B., Qvarfordt, P. and Chen, F. 2016. Tweetviz: Visualizing Tweets for Business Intelligence. In: *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*, July 2016, pp. 1153-1156, ACM.

Sleem-Amer, M., Bigorgne, I., Brizard, S., Dos Santos, L.D.P., El Bouhairi, Y., Goujon, B. and Varga, L. 2012. Intelligent semantic search engines for opinion and sentiment mining. *Next Generation Search Engines: Advanced Models for Information Retrieval,* pp. 191-215, IGI Global.

Tayouri, D. 2015. *The Human Factor in the Social Media Security–Combining Education and Technology to Reduce Social Engineering Risks and Damages.* Procedia Manufacturing, 3, 1096-1100.

Tziralis, G., Vagenas, G., & Ponis, S. 2009. Prediction markets, an emerging Web 2.0 business model: towards the competitive intelligent enterprise. In Web 2.0 (pp. 1-21). Springer, Boston, MA.

Wen, C., Teng, Z., Chen, J., Wu, Y., Gong, R. and Pu, J. 2016. socialRadius: Visual Exploration of User Check-in Behavior Based on Social Media Data. In: *Proceedings of the International Conference on Cooperative Design*, October 2016, Visualization and Engineering, pp. 300-308, Springer International Publishing.

Wongthongtham, P., & Abu-Salih, B. 2015. Ontology and trust based data warehouse in new generation of business intelligence: State-of-the-art, challenges, and opportunities. In Industrial Informatics (INDIN), 2015 IEEE 13th International Conference on (pp. 476-483). IEEE.

Wu, Y., Liu, S., Yan, K., Liu, M. and Wu, F. 2014. Opinionflow: Visual analysis of opinion diffusion on social media. *IEEE Transactions on Visualization and Computer Graphics*, 20(12), 1763-1772.

Yang, C. S. and Shih, H. P. 2012. A Rule-Based Approach for Effective Sentiment Analysis. *PACIS*, p. 181).

Yang, C.S. and Chang, P.C. 2015. Mining Social Media for Enhancing Personalized Document Clustering. In: *Proceedings of the International Conference on HCI in Business*, pp. 185-196, Springer International Publishing.

Yang, C.S. and Chen, L.C. 2014. Personalized Recommendation in Social Media: a Profile Expansion Approach. *PACIS*, p. 68.

Zeng, D., Chen, H., Lusch, R. and Li, S.H. 2010. Social media analytics and intelligence. *IEEE Intelligent Systems*, 25(6), 13-16.

Zhang, Z., Guo, C. and Goes, P. 2013. Product comparison networks for competitive analysis of online word-of-mouth. *ACM Transactions on Management Information Systems (TMIS),* 3(4), 20.

Zhang, Z., Li, X. and Chen, Y. 2012. Deciphering word-of-mouth in social media: Text-based metrics of consumer reviews. *ACM Transactions on Management Information Systems (TMIS),* 3(1), 5.

Zimmerman, C., & Vatrapu, R. 2015. The Social Newsroom: Visual Analytics for Social Business Intelligence. In: *Proceedings of the International Conference on Design Science Research in Information Systems*, pp. 386-390, Springer International Publishing.

Zimmerman, C.J., Wessels, H.T. and Vatrapu, R. 2015. Building a social newsroom: Visual analytics for social business intelligence. In: *Proceedings of the IEEE 19th International Conference*, Enterprise Distributed Object Computing Workshop (EDOCW), pp. 160-163, IEEE.

Journal of Intelligence Studies in Business
Vol. 8, No. 2 (2018) pp. 43-61
Open Access: Freely available at: https://ojs.hh.se/

Investigating the competitive intelligence practices of Peruvian fresh grapes exporters

Christophe Bisson[a,b] and Maria Mercedes Tang Tong[c]

[a] *DICEN, Site Val d'Europe Cité Descartes 5 Bd Descartes 77454 Marne-la-Vallée Cedex 2;*
[b] *ESSEC, 3 Avenue Bernard Hirsch, 95021 Cergy-Pontoise, France;*
[c] *Universidad de Lima, Peru*

Corresponding author (*): christophe.bisson@essec.edu

Received 17 May 2018 Accepted 15 August 2018

ABSTRACT This paper reports an empirical study of Peruvian fresh grapes exporters with the aim of delineating the behavioral and operational typology of their competitive intelligence practices. Cluster analysis was used as an exploratory tool to determine the correlation, if any, between the size of the company, grapes exports share of total exports, the percentage of the Red Globe variety in total grapes exports and the size of the grape farm with the typology and the average price received at export between August 2016 and July 2017. The behavioral and operational typology of competitive intelligence practices model, developed by Wright *et al,* (2012), was used. The findings reveal that exporters have a positive behavior towards competitive intelligence practices, but cannot make good use of them due to a lack of knowledge, and deficiencies in organization and in technological and IT systems support. As 37 companies participated in this experiment, this study could be extended to all non-traditional Peruvian agricultural exports. It has been possible to identify areas where changes are needed to enable these exporters to perform at a higher level of competence. In addition, it appeared that a slightly higher level of attitude and IT systems support pays off as medium-sized companies achieved a higher price per ton compared to big companies. This study is the first to present a typology of competitive intelligence practices in Peru and is one of the very first to study competitive intelligence in this country and agriculture.

KEYWORDS Behavior, competitive intelligence, grapes exporters, Peru, typology

---

1. INTRODUCTION

As companies face fiercer competition and a more uncertain environments, competitive intelligence (CI) is gaining ground (Blenkhorn and Fleisher 2005; Bisson and Yasar Diner, 2017). The Global Intelligence Alliance (GIA), using data from surveys done on the same sample in 2009 then in 2011, reported that the percentage of companies integrating CI functions increased from 63% to 76% in this period (GIA, 2011) and that decision making was 15% more efficient in companies that utilize CI functions (GIA, 2013).

CI originated from military intelligence and dates back to Sun Tzu and is thereby an art in addition to being a science (Prescott, 1999). Its systematic use in the commercial and business world is fairly recent and many academics have studied their country's CI practices (Calof *et al.,* 2015). Soilen (2013) reviewed fifty-one articles written by eighty-three authors, mostly from the United States, Canada and the United Kingdom, published in the *Journal of*

Competitive Intelligence and Management (JCIM) between 2004 and 2008. He found that the main topics of research were the development of CI in general or in specific countries, followed by studies defining CI and studying its growth in time, and finally, business intelligence and its applications.

Little research has been conducted on the application of CI in developing countries (Ifan *et al.*, 2004; Zhan and Chen, 2009; Wright *et al.*, 2013; Du Toit, 2013; Du Toit and Sewdass, 2014; Rodriguez Salvador and Salinas Casanova, 2012; Rodrigues and Thome e Castro, 2017) with only a few isolated efforts focusing on the Spanish speaking communities of South America (Aguirre, 2015; Guarrochena and Paul, 2013; Salazar *et al.,* 2014; Villaroelg *et al.,*2015).

Thus, this study explores for the first time the CI practices of export companies in the fresh grape sector in Peru. Hence, it could inspire Peruvian companies and promote more studies of CI in South America. In addition, very few studies about CI in agriculture have been undertaken (Bisson 2014). The purpose of this study is to create a typology of Peruvian fresh grapes exporters' CI practices and to investigate the relationship between the size of the company, the share of grapes exports in total exports, the percentage of the Red Globe variety in total grapes exports and the size of the farm with their CI practice levels and the average price received at export.

The remainder of the paper is organized as follows: we first provide a brief comparison of the conception of competitive intelligence in English and Spanish literature, then we deal with CI in Peru followed by the importance of fresh grapes in Peruvian exports. The methodology used in this research is described and the results are then presented and discussed. Finally, we conclude with an examination of the implications and limitations of this research and suggest further research that may be undertaken.

2. THEORETICAL BACKGROUND

2.1 The competitive intelligence conception in English and Spanish literatures

In the English literature, there is no universal definition of CI accepted by all (Du Toit, 2015; Wright et al., 2009). Haddadi et al. (2010) emphasize that the lack of an accepted definition renders this field unstable. CI was developed in the early 1980s (Presscot, 1999) in the US, focusing originally on competitors under the influence of Porter (1980) and was then broadened to include all actors in the market. Although it is commonly accepted that CI makes use of information from outside the organization (and is thereby based on monitoring or scanning the organization's environment), some authors (e.g. Wright, 2011) consider that CI should also encompass internal information to fulfil the needs of decision makers.

Calof et al. (2015) categorize the definitions by those who focus on the objectives of CI, i.e. to enlighten decision makers and those who explain it by how CI is performed thereby centered on the intelligence cycle. This cycle has four steps (Kahaner, 1997): i) planning and direction; ii) collection; iii) analysis; and iv) dissemination. Thus, after defining the key intelligence topics, information is gathered, analyzed and the results are disseminated to people who triggered the cycle. Pellissier and Nenzhelele (2013) studied 50 CI definitions and determined that 38 referred to CI as a process and 4 as a product. In terms of its objectives, CI has been defined by Du Toit (2013, 30) as "… a strategic tool to facilitate the identification of potential opportunities and threats". In the same vein, Presscott and Miller (2001) define it as any actionable intelligence that could provide a competitive edge. As a process, Kahaner (1998, p.16) states that "Competitive intelligence is a systematic program for gathering and analyzing information about your competitors' activities and general business trends to further your own company's goals". Likewise, Fleisher (2004, 56) defines it as a "… systematic process by which organizations ethically gather and analyze actionable information about competitors and the competitive environment and, ideally, apply it to their decision-making and planning processes to improve their performance". In contrast, Rouach and Santi (2001, p.553) suggest it is a creative process, or "the art of collecting, processing and storing information to be made available to people at all levels of the firm to help shape its future and protect it against current competitive threats: it should be legal and respect codes of ethics; it involves a transfer of knowledge from the environment to the organization within established rules".

Soilen (2016) argue that definitions of CI and marketing intelligence are quite similar and overlapping, addressing the same phenomenon, but studied by different academic disciplines. Du Toit (2015), based on 338 published peer-reviewed articles from 1994 to 2014 in the ABI/Inform database, found that the most popular term used in the literature is CI, followed by business intelligence and marketing intelligence.

Compared to the English literature, the main difference in the Spanish literature is that competitive intelligence is linked to the term 'technological watch' in accordance with the norm UNE 166006:2011 (the Spanish Association for Standardization and Certification [Aenor 2018]) and has risen in the Spanish speaking community independently to the English speaking CI and marketing intelligence academic communities. For instance, Professor Escorsa has written numerous articles in Spanish about technological watch while dealing with CI (see, for example, Escorsa and Maspons, 2001). In a similar vein Rodriguez Salvador and Slinas Casanova (2012) suggest that the ultimate objective of CI is to support innovation.

2.2 Competitive Intelligence in Peru

In Peru, based on the largest number of publications found by the search engine of the Peruvian repository for theses and academic papers, the most common terms associated with CI are business intelligence followed by marketing intelligence (Concytec, 2018). From the total of 375 titles that appear in a search carried out on March 27th 2018, all were monographs or news items and there were only nine peer-reviewed articles, from which only two are related to the research topic. These two articles are a study that covers ten in-depth interviews about factors needed to promote foresight and competitive intelligence in 2040 (Inche Mitma et al., 2016) and a survey of 28 Peruvian exporters and importers about implementations of market intelligence programs in their companies (Tang Tong, 2015).

The lack of peer-reviewed articles about intelligence processes or programs to scan the environment in order to be more competitive in Peru reflects the poor efforts to promote CI as well as the lack of human resources needed to develop CI as stated in the report by the National Council of Science, Technology and Technological Innovation (Concytec, 2017): i) there are only two Public Institutes of Research which have technology transfer units and that perform activities of technological surveillance; one of these is the Peruvian Technological Institute of Production (ITP). ITP has been recognized as the first organization in Latin America to obtain a certification for technological watch and CI according to the Norm UNE 166006:2011 (Aenor, 2018); ii) only two companies are offering this service of technological surveillance in the domestic market; iii) there are very limited educational offerings at universities and institutes. Recently, Concytec launched a five year-program (2017-2021) to promote capabilities in technological watch and CI as a means to achieve higher innovation, following the successful experiences observed in Argentina and Colombia. Indeed, these two countries have set up technological observatories, providing access to scientific, technological and competitive knowledge that can be adopted nationally (Concytec, 2017).

Some efforts to help exporters have been made through the Peruvian Export and Tourism Promotion Agency (Promperu), providing research studies of main export markets, which were developed by the market intelligence unit and are available on their web site (Promperu, 2018). However, there are no reports monitoring the main markets.

2.3 The importance of fresh grapes in Peruvian exports

Since the beginning of the 21st century, Peru has emerged as one of the fastest-growing and most stable economies in Latin America, with an average annual growth rate of 5.1% between 2007 and 2016 (the Central Reserve Bank of Peru [BCRP], 2018; World Bank Group, 2017). Non-traditional agricultural exports, with fresh grapes making up the largest share, have shown an impressive compound annual growth rate (CAGR) of 13.4% in the same period, accounting for 13% of total exports in 2016 (the National Superintendence of Customs and Tax Administration [Sunat], 2017).

The Peruvian fresh grapes exports sector has been developing since the end of the 1990s and has grown at double digit rates driven by private investments and modern technologies, and the sector is vertically integrated and created with the sole purpose of serving the exports market (Meade et al., 2010; World Bank Group, 2017). As a result, Peru is the world's fifth largest exporter of fresh grapes, accounting for 6.3% of worldwide grape exports in 2016 (International Trade Center [ITC], 2017).

The Ministry of Agriculture and Irrigation of Peru (Minagri, 2017) estimated that the total production was 689,800 metric tons (MT) in

2016. This number has more than doubled since 2010, as a consequence of a wider growing area. The last census in 2012 estimated there to be about 43,800 hectares dedicated to grapes, covering both wine production and fresh grapes for consumption (the National Institute of Statistics and Information [INEI], 2013). This figure is likely to have also increased and it is estimated that there are 30,000 hectares in Peru dedicated to fresh grapes, where the Red Globe variety is the most common with 80% of the total production (Fernandez-Stark et al., 2016).

The increase in growing areas is mainly due to the perfect match between the Peruvian production months and the months of lower production in the northern hemisphere. Almost half of the production is exported during the higher production season i.e. from August to April, when the export price is on average three times higher compared to the local price (BCRP, 2018; Minagri, 2017; Sunat, 2018).

As more companies got involved in exporting grapes due to higher prices, Peruvian exports grew rapidly with a CAGR of 24.2% between 2010 and 2016, impacting the world supply and leading to lower prices in recent years (ITC, 2010-2017).

3. METHODOLOGY

3.1 Sample and procedure

For the purposes of this study the model developed by Wright et al. (2012) is used, a behavioral and operational typology of CI practice applied to SMEs and construed as being robust (Ross et al., 2012; Gaspareniene et al., 2013; Smith, 2012; Bisson, 2013; Toker et al., 2016). This model was itself adapted from the study of Wright et al. (2002) of CI active firms in the UK which addressed four strands: attitude, gathering, use and location. This model has inspired further work and replication studies carried out by Adidam et al. (2009), April and Bessa (2006), Bouthillier and Jin (2005), Dishman and Calof (2008), Liu and Wang (2008), Oerlemans et al. (2005), Priporas et al. (2005), Rodrigues and Thome e Castro (2017) and Wright et al. (2009). Wright et al. (2012) added two new strands: technological support ("as degree of investment made to assist with gathering competitive information") and IT support ("as the type of systems used to manage the flow of competitive information"). In this way each strand is related to specific questions that later can be translated into a typology verdict for each exporter.

A questionnaire using both closed and open questions was used to gather the data set. Self-declared position statements were also included in the questionnaire to either confirm or contradict answers given within each section. The latter served as a clarification mechanism to identify any contradiction in a typology verdict.

The questionnaire was available on-line in Spanish and a secured link was created for each exporter. The target group was the Peruvian grape exporters that had exported grapes according to the harmonized tariff code 08.06.10.00.00 in 2016 available in Sunat (2017). Peruvian customs provided a list of exporters that was then cleaned for the purposes of this research. The eligible sample comprised 80 export companies.

All companies were contacted by telephone and/or reached by e-mail to be invited to take part in this study between October 2017 and March 2018. A total of 37 questionnaires were completed. The sample used in this research represents more than 60% of the total exporters (detailed in Table 1). The unit price achieved by the companies of the sample was higher than the average for all companies. Companies were classified as being a big, medium, small or micro company using as a reference the European Union definition of an SME in terms of turnover and employee numbers (EU Commission, 2003).

Table 1

Characteristics of the sample

UNIVERSE season 2016/17						SAMPLE season 2016/17					
Size	No. of Companies		Total exports$		Exports $/MT	Size	No. of Companies		Total exports$		Exports $/MT
Big	14	8%	280,716,249	40%	2,267	Big	8	22%	187,290,212	44%	2,336
Medium	26	15%	191,600,688	28%	2,404	Medium	15	41%	167,838,700	39%	2,518
Small	51	30%	175,494,106	25%	2,080	Small	10	27%	60,790,851	14%	2,309
Micro	81	47%	46,430,922	7%	1,922	Micro	4	11%	9,639,173	2%	1,988
TOTAL	172	100%	694,241,965	100%	2,225	TOTAL	37	100%	425,558,936	100%	2,391

- The season starts in August and finishes in July the following year.

More than half of the interviewees were top management, holding positions of CEO or Chairman of the Board, one fourth were management positions reporting to the CEO, and the remaining respondents were those reporting to first line management.

Most companies stated that they exported more than 75% of their sales and 32 out of 37 companies were vertically integrated throughout the main steps of cultivation, harvesting, processing and export. Five of the companies did not cultivate grapes but acted as processors and exporters on behalf of other producers.

The size of the farm was asked to those involved in cultivation and most companies stated they had more than 100 hectares for grapes cultivation. According to the last farm structure survey carried out in the European Union in 2013, the largest agricultural holding size was found to be more than 100 hectares and these made up 2.7% of 12 million farms accounting for over 30% of standard output across the EU (European Commission, 2013). Similarly, in the latest Peruvian Agriculture Census carried out in 2012, the largest farms were also found to be larger than 100 hectares and they were estimated to be 0.9% of 2.2 million farms (INEI, 2013).

3.2 Analytical approach

The same set of descriptors utilized by Wright et al. (2012) was used (see Appendix 1), and the findings from this study were applied to this behavioral and operational typology of CI to reach verdicts regarding levels of gathering, attitude, use, location, IT systems and technology support. Furthermore, cluster analysis was used as an exploratory tool (Kaufman and Rousseeuw, 2005) to investigate whether there was any correlation between the size of the company, grapes exports share of total exports, the percentage of the Red Globe variety in total grapes exports and the size of the grape farm with the typology and the average price received at export between August 2016 and July 2017.

4. RESULTS AND DISCUSSION

4.1 Gathering

This section asked about the type of information they collected, the sources they used, how much competitive information they obtained from their own employees, how they prepared their employees to address competitors, what type of financial return they expected from their CI effort and how much financial support was provided for CI activities.

With regards to the type of information they collected, 284 responses were recorded, with customers, competitors, products in their market, suppliers and scientific articles and publications taking the top five places, closely followed by job market, laws, economy, politics and taxation policies. The items that were revealed as being of less interest were ISO standards, patents, industrial processes, social and finance. Interestingly, only one respondent included weather information, which is of utmost importance in agriculture, another respondent included certification requirements, which are compulsory for this kind of business due to food safety and traceability issues, and another respondent included yields in other countries, and phytosanitary barriers among non-tariff as well as tariff trade barriers.

The most popular source of information was stated to be trade fairs followed by industry experts and industry magazines. This is indicative of reliance on a well-informed set of sources. An additional source of information was input received from employees, as 86% of respondents stated that they obtained either a moderate or high amount of competitive information from their own employees. However, the most sophisticated sources such as written evidence from verified sources, competitor research obtained from an external source, media analysis, management consultants and forecasting models were the least used.

About 70% of respondents stated that they always or often trained and prepared their employees before they went to trade shows, exhibitions, conventions and other public events to make them aware of the type of information they should look for. However, the remaining 30% did this only ‘occasionally’ or ‘never’. Only 59% of respondents said that they always or often briefed their employees on what they should not talk about to competitors, which demonstrates that companies are paying less attention to this area. This leaves 41% who are either naive or reckless about the importance of protecting the company’s sensitive information.

Considering that 81% of respondents stated that they evaluated the reliability of their sources of information, it is interesting to note that this task is not an easy one as the top three barriers to effective competitive information gathering in the open question section, were reported as: i) access to the information; ii) reliability of the information; iii) lack of resources (mostly time) which were indicated by 57%, 54% and 38% of the respondents, respectively.

Concerning the financial support given by the organization for the task of monitoring the competitive environment, about 57% of respondents considered the support given to be adequate to do a reasonable job or enough to do a good job. On the other hand, 30% stated that: i) no funds were available as the tasks were done by interested people rather than intelligence experts; ii) funds were provided if an immediate financial benefit could be produced; iii) minimal support was provided to cover the basic tasks and simple gathering. The remainder stated that the activity received a set budget or that funds were available on request.

Based on the provided answers, the overall verdict inclined towards a hunter gathering level. However, the self-declared control statement showed that the verdict may be more nuanced as half of the companies used only public domain sources for their competitive information. Thus, the verdict is hunter gatherer, but several of these companies take their desire for real as they are not using sophisticated ways to collect information.

4.2 Attitude

Regarding how often the firm collected information about competitors, technologies and customers, the most frequent answer was weekly for customers and competitors while both monthly and irregularly were answered 'when it becomes available or required for a project' for technology. Even though there seems to be a regular process to gather data, 41% agreed that it is not an organized process, and only 5% of the companies had a written process and a system dedicated to CI. Therefore 95% of firms have no formalized process or dedicated system to handle gathered information.

Furthermore, 11% claimed that their companies provided 'full commitment for understanding competitors' and 70% stated that there was either 'active support for current activities' or 'just about sufficient for immediate needs'. These findings are in line with the self-declared control statement in which 30% 'try to understand specific questions for one-off projects', 41% 'try to understand the market in the short term' and 22% had an integrated competitive information process where competitors were monitored to anticipate their moves and to plan a reaction. Only 8% agreed that 'we are too busy thinking about today to worry about tomorrow'. Here the verdict was a task-driven attitude but significantly biased towards both an immune and operational stance.

4.3 Use

When asked how they used the collected information, 68% of respondents stated that they use it for both short and long-term decision making and 54% for scenario planning, leading to a verdict of strategic user. However, 41% stated that 'there is no organized process for feeding CI output into the decision-making processes, leading to a verdict of Joneses user.

Concerning the impact different factors have in the company decision making, 'customer demands' was the most frequent choice, followed by 'competitors' long term predicted behavior', 'competitors' short term predicted behavior' and 'technological/technology standard changes'. These are congruent with the self-declared control statement in which 38% 'use competitive information to help make decisions about price changes and promotional efforts' and 46% use competitive information to identify opportunities and threats as well as to build scenarios. These findings suggest a verdict of strategic user but with a strong tendancy towards a Joneses user stance.

4.4 Location

In this section, participants were asked whether employees knew who to pass information on to when they acquired it, and 92% of respondents stated either 'always' or 'often', with only 8% stating they knew 'occasionally'. The top four departments that took responsibility for collecting CI were first sales (59%) and then general management (43%), followed by manufacturing & production (27%) and research & development (22%) with 22% of respondents stating also that all departments take responsibility. The latter response suggests that some companies work in a loose manner as they do not have a clear idea of who should take overall responsibility.

When asked whether a dedicated intelligence unit is essential to successfully accomplish the monitoring task, only 16% responded with 'always' and 30% 'sometimes' while 38% stated this to be 'a good idea but not always essential'. The remainder responded with either 'not needed at all' or 'it seems to work well without a dedicated unit'.

Based on the above findings, it came as no surprise that 89% of respondents stated that they did not have a dedicated intelligence unit, although 54% did have a person in their firm whose job is to gather, analyze, disseminate and store the competitive information, and in 65% of the cases this person participated in senior management meetings. In sum, the verdict was an ad-hoc location approach.

4.5 Technology support

This strand deals with the type of tools used by the companies to gather information. The most frequently used tools were websites (92%) and Google (86%), followed much less frequently by specialized databases such as Derwent, Dun & Bradstreet and Euromonitor (41%) and specialized websites, for example Espacenet for patents (22%). This is in line with the self-declared control statement in which 72% of respondents stated they 'use common, freely available tools for web searching, such as Google'. However, 14% of respondents 'use full versions of meta-search engines and are also familiar with specialist databases for patent and financial information' and 14% 'use software that allow users to collect, analyze and disseminate information automatically'. The verdict was overwhelmingly a simple technology support stance.

4.6 IT Systems

This section addresses the IT systems used to manage competitive information in the companies. About 49% of the respondents stated that they did not use any systems at all to manage their competitive information and agreed with the statement that 'it is in our minds and we rely on our memories'. This contrasted with the next largest categories, chosen to a much lesser extent, with 16% stating that 'we use IT systems to manage competitive information but to ensure the safety of our information we prefer paper records and do not really like relying on computers, or somebody else', 19% stating they used off-the-shelf and 14% stating they used a bespoke development.

This is in line with the control self-declaration in which 38% agreed they did not use IT systems to manage competitive information and 'rely on our memories and the good will of staff to share what they learn' and 22% stating they 'prefer to stick to traditional methods of managing competitive information by using paper records' and agreed with the statement that they 'do not really trust computers'. However, this is in contrast to the 22% which claimed to have designed their own in-house system unique to the firm and its needs. Here the verdict was a dismissive IT systems stance with a strong tendency towards bespoke IT systems.

4.7 The typology of Peruvian grape exporters' CI practice levels

The verdicts for each strand i.e. gathering, attitude, use, location, IT systems and technology support are summarized in Figure 1. The Peruvian grape exporters appear to be aware of the importance of CI but they lack knowledge, organization and dedicated IT. Hence, thanks to the evaluation carried out in this study, companies can see the path to follow that should lead them towards higher CI practice levels to help them better address a faster and harsher competitive environment.

4.8 Cluster analysis by size of company

With regards to the six strands of the CI typology studied, practices among big, medium, and small & micro companies are rather similar to the findings for the total sample as shown in Table 2 (for more details, see Appendix 2). However, big companies have a more immune attitude compared to the task driven attitude of medium companies and the operational attitude of small companies. Furthermore, about the use of information, if big and small & micro companies are at a strategic level, medium companies are the lowest one.

In general, for all the CI strands, the percentage of small & micro companies are at higher levels. One can construe that these small & micro companies need to be more aggressive to survive as they compete with bigger companies and that consequently they seem to be more aware of the value of information for competitiveness.

Despite this, medium companies registered higher average prices (Free On Board [FOB] Peruvian port US$ 2,518 per metric ton) compared to big companies (FOB US$ 2,336 per metric ton). The small & micro companies registered the lowest average price (FOB US$ 2,259 per metric ton). This cannot be interpreted to mean that big companies have a more

US 2,677 per metric ton)

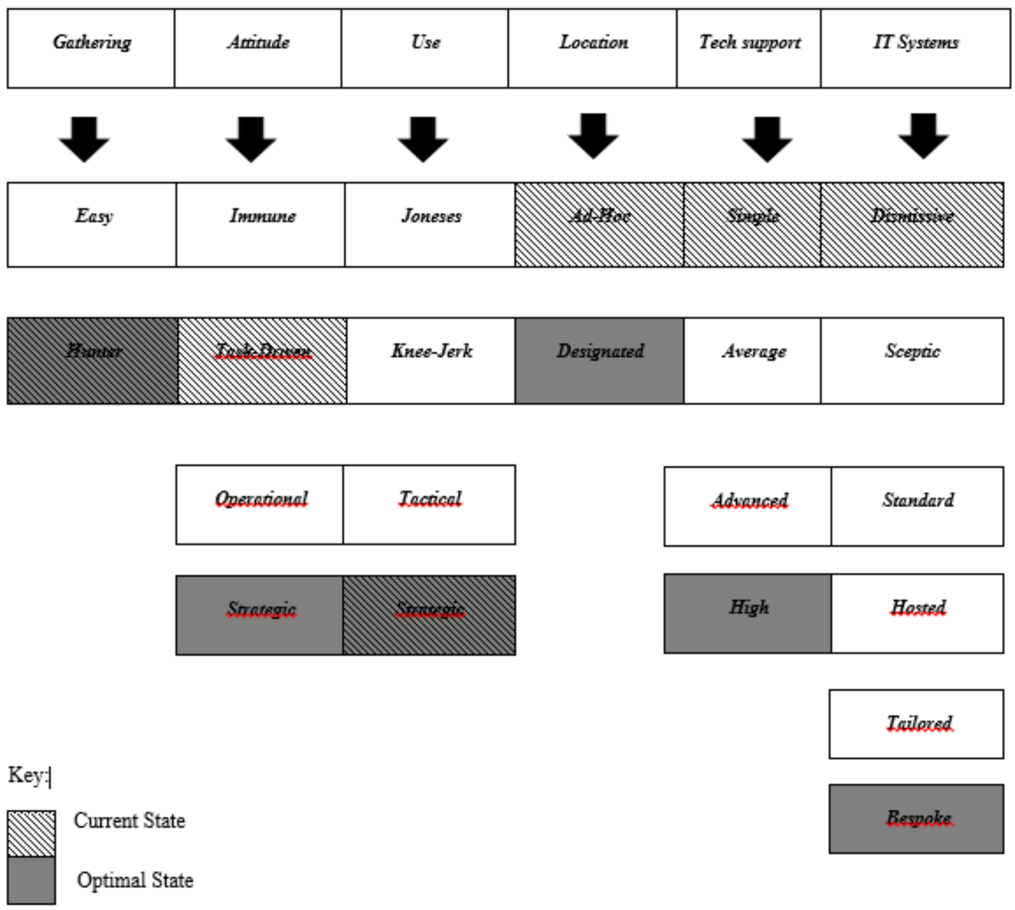


Figure 1 The behavioral and operational diagnostic typology of Peruvian grape exporters' competitive intelligence practice.

challenging job placing their grapes in the market compared to medium sized companies as it is shown later that the larger the grape farm the better results in price per ton. This suggests that a positive behavior towards CI pays off as medium sized companies show a higher level in this strand compared to big companies, with more cases of technology and IT support being utilized. This also suggests that the CI level is independent of the size of the company in line with the results of Priporas et al. (2005).

4.9 Cluster analysis by percentage of grapes exports in total exports

This cluster confirms that those companies that do not concentrate primarily on grapes, with grapes representing less than 75% of their total exports, have a stronger attitude towards an operational stance compared to those which are less diversified and tend towards a task-driven attitude. However, it shows that a concentration as opposed to a diversification strategy pays off as the price per ton is significantly higher in those companies concentrating on grapes (FOB Peruvian port compared to those that do not (FOB US$ 2,133 per metric ton).

4.10 Cluster analysis by percentage of Red Globe variety in total grapes exports

Companies with a concentration of the Red Globe variety higher than 50% received a significantly lower price (FOB Peruvian port US$1,881 per metric ton) compared to those that have less concentration in this variety (FOB US$ 2,605 per metric ton). However, this cluster shows homogeneous results compared to the sample. It is worth noticing that higher value grapes increase the labor and handling costs, which moderate the variety choice (Fernandez-Stark et al. 2016).

4.11 Cluster analysis by size of farm

This cluster was analyzed based on those companies that have grape cultivation. It indicates that the companies with less than 100 hectares and more than 501 hectares behave differently from the average sample. Indeed, those companies with less than 100 hectares show a stronger attitude towards an operational stance, which somehow is

Table 2 Cluster analyses.

Strand	Gathering	Attitude	Technology	Information System	Use	Location
Cluster	Verdict	Verdict	Verdict	Verdict	Verdict	Verdict
Company size*						
Big	G2	A1	TS1	IS1	U4	L1
Medium	G2	A2	TS1	IS1	U1	L1
Small & micro	G2	A3	TS1	IS1	U4	L1
% of grapes exports in total exports**						
Higher than 75%	G2	A2	TS1	IS1	U4	L1
Lower than 75%	G2	A3	TS1	IS1	U4	L1
% of red globe in total grapes**						
Lower than 50%	G2	A2	TS1	IS1	U4	L1
Higher than 50%	G2	A2	TS1	IS1	U4	L1
TOTAL	G2	A2	TS1	IS1	U4	L1
Grapes farm size						
Lower than 100 hectares	G2	A3	TS1	IS1	U4	L1
between 101 and 500 hectares	G2	A2	TS1	IS1	U4	L1
Higher than 501 hectares	G2	A3	TS1	IS1	U1	L1
TOTAL	G2	A2	TS1	IS1	U4	L1

Sources:
* - Peru: Top Publications (2018)
** - Sunat (2016-2017)

translated into a higher level of IT systems use, and the use of the information strategically. On the other hand, those companies with more than 501 hectares also show a stronger attitude towards an operational stance, which is also translated into different levels of higher IT systems with more technology support, but they do not use the information strategically.

This cluster confirms that the largest grape farms, with more than 501 hectares, obtained a better price (FOB Peruvian port US$ 2,444 per metric ton) compared to the lower prices seen for 101-500 hectare grape farms (FOB US$ 2,413 per metric ton) and much higher prices than 100 hectares grape farms (FOB US$ 1,932 per metric ton). This can be interpreted to indicate that there is an advantage in having a higher critical mass volume for exports, since some importers prefer larger volumes from a few growers that can ensure quality consistency, food safety and traceability.

5. CONCLUSION

This paper aims to create a typology of Peruvian fresh grapes exporter CI practices. Overall, this sector shows positive behaviors towards CI but cannot make the most of it due to the lack of technological and IT systems support, lack of knowledge and dedicated organizational structures. The first verdict is that this sector displays the hunter gathering stance, which is a key indicator to engage in CI practice. However, evidence also suggests that there is still too much effort spent on easy gathering from public sources producing volume, not value. The second verdict is that exporters show a task-driven attitude where questions are asked and answered with little value added. In order to reach the ideal state of a strategic attitude, top management should embrace CI as essential for future success, addressing ‘what if’ questions for both short and long-term decisions, anticipating changes and planning possible courses of action. The third verdict is that this sector is a strategic user, which is the optimum state but is strongly biased towards Joneses user as the knowledge learnt is not retained for the future. The fourth verdict is ad-hoc location instead of dedicated location for CI practice, despite the fact that almost half of respondents have a person who gathers, analyzes, disseminates and stores competitive information. In order to have a successful CI program, it is necessary to define roles and responsibilities with a specific location within the organization. This way redundant work is avoided and it empowers the person in charge to develop technical and cognitive skills to deliver the right CI to the right person at the right time. The fifth verdict is that this sector uses very simple tech support, which does not require specific knowledge, commonly using spreadsheets for their analysis and accessing web sites displaying old information that provides limited value. With globalization, increasing data complexity and speed of change, it is of the utmost importance to invest in integrated systems (e.g. scanning systems) that provide information in real time and allow this information to be aggregated. The last verdict is dismissive IT systems support as companies do not use any IT systems to manage strategic information.

The second aim of this paper explored whether the size of the company or the export level of these companies impact their CI practice level. According to the cluster analysis by size of company, CI practice level is independent of the size of the company as big, medium and small & micro companies show almost homogeneous results among the six strands. However, it seems that a slightly higher level of attitude and IT systems support pays off as medium companies show a higher price per ton compared to big companies. This does not mean that large companies have to struggle more to place more volume as cluster analysis by size of farm makes it clear that the larger the grapes farm size the higher the price per ton. The cluster analysis of grapes exports in total exports suggests there are advantages to specialization instead of diversification, as companies with grapes exports representing more than 75% of their total exports receive a higher price per ton compared to those whose grapes exports were below 75% of their total exports. Finally, the cluster analysis of the ratio of the Red Globe variety in total grapes exports, shows that significantly lower prices are received by companies that have more than 50% Red Globe in their total grapes exports. However, this cluster shows homogeneous results compared to the sample.

The results of this study provide empirical evidence to the Peruvian Government authorities about the need to promote training and the adoption of dedicated technology among companies in order to achieve higher levels of CI practices. Furthermore, Peruvian authorities as well as other South American governments can benefit from the experience of other countries that have government sponsored CI programs, specifically Canada (Brouard, 2006; Tanev and Bailetti, 2008; Tarraf and Molz, 2006), France (Bisson, 2010, 2013; Salles, 2006; Smith et al., 2010) and Switzerland (Begin et al., 2007).

5.1 Limitations and further research

As the sample size is limited, this experiment could be extended, for example, to all non-traditional agricultural Peruvian exports to confirm the findings reached in this study and to be able to address SMEs, which are known as PYMES in Latin America, to help Peruvian authorities to better address their needs.

Based on the experiences in Canada, evaluations of CI programs do not measure the direct economic impact and Calof (2017) points out that this needs to be addressed in future research. Therefore, it could be quite interesting to create a longitudinal analysis of this non-traditional agricultural Peruvian exports sector to measure the impact of CI workshops, training and dedicated IT tools on the competitive and financial performances of these companies.

6. REFERENCES

Adidam, P., Gajre, S. and Kejriwal, S. (2009). Cross-cultural competitive intelligence strategies. *Marketing Intelligence and Planning*, 27(5), 666-680.

Aguirre, J. (2015). Strategic intelligence: a system to manage innovation. *Estudios Gerenciales*, 31(134), 100-110.

April, K. and Bessa, J. (2006). A critique of the strategic competitive intelligence process within a global energy multinational. *Problems and Perspectives in Management,* 4(2), 86-99.

Asociacion Española para la Normalizacion y Certificacion (2011). *Norm UNE 166006:2011*. Retrieved from: http://www.aenor.es/aenor/normas/normas/fichanorma.asp?tipo=N&codigo=N0046930#.Wpmn1JNuai4

Asociacion Española para la Normalizacion y Certificacion (2016). *AENOR Perú otorga certificado de vigilancia tecnológica a ITP*. Retrieved from: http://www.aenor.es/aenor/actualidad/actualidad/noticias.asp?campo=1&codigo=43184&tipon=1#.Wpmi8JNuai6

Banco Central de Reserva del Perú (2017). *Exchange rate*. Retrieved from:

https://estadisticas.bcrp.gob.pe/estadisticas/series/

Begin, L., Deschamps, J. and Madinier, H. (2007). *An interdisciplinary Approach of Competitive Intelligence*. Cahiers de recherche du CRAG, No. HES-SO/HEG-GE/C-07/4/1-CH.

Bisson, C. (2010). Development of competitive intelligence tools and methodology in a French hightech SME. *Competitive Intelligence Magazine,* 13, 18-24.

Bisson, C. (2013). *Guide de Gestion Stratégique de l'information pour les PME*. Montmoreau: Les2encres.

Bisson, C. (2014). Exploring competitive intelligence practices of French local public agricultural organisations. *Journal of Intelligence Studies in Business*, 4(2), 5-29.

Bisson, C. and Yasar Diner, O. (2017). Strategic Early Warning System for the French milk market: A graph theoretical approach to foresee volatility. *Futures*, 87, 10-23.

Blenkhorn, D. and Fleisher, C. (2005). *Competitive intelligence and global business.* Westport: Praeger.

Bouthillier, F. and Jin, T. (2005). Competitive intelligence and webometrics. *Journal of Competitive Intelligence and Management*, 3(3), 19-39.

Brouard, F. (2006). Development of an expert system on environmental scanning practices in SME: Tools as a research program. *Journal of Competitive Intelligence and Management*, 3 (4), 37-55.

Calof, J. (2017). Reflections on the Canadian Government in competitive intelligence - programs and impacts. *Foresight*, 19(1), 31-47.

Calof, J., Richards, G. and Smith, J. (2015). Foresight, Competitive Intelligence and Business Analytics - Tools for Making Industrial Programmes more efficient. *Foresight-Russia*, 9(1), 68-81.

Comisión de Promoción del Perú para la Exportación y el Turismo (2018). *Inteligencia de mercados.* Retrieved from: http://www.siicex.gob.pe/

Consejo Nacional de Ciencia, Tecnología e Innovación Tecnológica (2017). *Programa Especial de Prospectiva y Vigilancia Tecnologica.* Retrieved from: https://portal.concytec.gob.pe/images/noticias/ProgramaVT_DocFinal_Vr12_ConsultaPublica.pdf

Consejo Nacional de Ciencia, Tecnología e Innovación Tecnológica (2018). *National digital repository of science, technology and innovation.* Retrieved from: https://alicia.concytec.gob.pe

Dishman, P. and Calof, J. (2008). Competitive intelligence: a multiphasic precedent to marketing strategy. *European Journal of Marketing,* 42(7/8), 766-785.

Du Toit, A. (2013). Comparative study of Competitive Intelligence Practices between two retail banks in Brazil and South Africa. *Journal of Intelligence Studies in Business,* 3(2), 30-39.

Du Toit, A. (2015). Competitive intelligence research: an investigation of trends in the literature. *Journal of Intelligence Studies in Business,* 5(2), 14-21.

Du Toit, A. and Sewdass, N. (2014). Competitive intelligence practices in Brazil: an exploratory study. *Mousaion,* 32(1), 84-97.

Escorsa, P. and Maspons, R. (2001). *De la vigilancia tecnológica a la inteligencia competitiva.* Madrid: Prentice-Hall.

European Commission (2003). Commission recomendation of 6 May 2003 concerning the definition of micro, small and medium-sized enterprises. *Official Journal of the European Union,* 124, 36-41.

European Commission (2013). *Structure and dynamics of EU farms: changes, trends and policy relevance.* Retrieved from: https://ec.europa.eu/agriculture/sites/agriculture/files/rural-area-economics/briefs/pdf/09_en.pdf

Fernandez-Stark, K., Bamber, P. and Gereffi, G. (2016). *Peru in the Table Grape Global Value Chain.* Retrieved from: http://www.cggc.duke.edu/pdfs/2016%20Duke%20CGGC%20Grape%20GVC%20Report%20Peru.pdf

Fleisher, C. S. (2004). Competitive Intelligence Education: Competencies, Sources, and Trends. *Information Management Journal,* 38(2), 56-62.

Gaspareniene, L., Remeikiene, R. and Gaidelys, V. (2013). The opportunities of the use of competitive intelligence in business: literature review. *Journal of Small Business and Entrepreneurship Development*, 1(2), 9-16.

Global Intelligence Alliance (2011). *Market intelligence in global organizations: survey* *findings in 2011*. Retrieved from: https://www.mbrain.com/wpcontent/uploads/2015/04/10852.pdf

Global Intelligence Alliance (2013). *The state of Market Intelligence in 2013: Global MI Survey Findings*. Retrieved from: http://www.biia.com/the-state-of-market-intelligence-in-2013-global-mi-survey-findings

Guarrochena, M. d. A. and Paul, L. M. (2013). Strategies of information management associates with competitive intelligence: appropriation practice in support exporters organizations. *Vision de Futuro,* 17(2), 168-185.

Haddadi, A., Dousset, B. and Berrada, I. (2010). *Xplor Everywhere — The competitive intelligence system for mobile*. Retrieved from: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5945635&isnumber=5945564

Ifan, H., Dou, J. M., Manullang, S. and Dou, H. (2004). Developing competitive technical intelligence in Indonesia. *Technovation*, 24 (12), 995-999.

Inche Mitma, J. L., Chung Pinzas, A. R. and Ruiz Lizama, E. (2016). Prospective and Competitive intelligence in 2040 to promote programs innovation in Peru. *Industrial Data,* 19(1), 38-44.

Instituto Nacional de Estadistica e Informatica (2013). *Resultados definitivos IV Censo Nacional Agropecuario 2012*. Retrieved from: http://proyectos.inei.gob.pe/web/DocumentosPublicos/ResultadosFinalesIVCENAGRO.pdf

International Trade Centre (2008-2017). *List of exporters of selected product 080610 Grapes, fresh*. Retrieved from: http://www.trademap.org

Kahaner, L. (1997). *Competitive Intelligence. How to gather, analyze and use information to move your business to the top*. New York: Touchstone.

Kahaner, L. (1998). *Competitive Intelligence.* New York: Touchstone.

Kaufman, L. and Rousseeuw, P. (2005). *Finding groups in data: an introduction to cluster analysis.* New York: Wiley-Interscience.

Liu, C. H. and Wang, C. C. (2008). Forecast competitor service strategy with service taxonomy and CI data. *European Journal of Marketing,* 42 (7/8), 746-765.

Meade, B., Baldwin, K. and Calvin, L. (2010). *Peru: an emerging exporter of fruits and vegetables*. Retrieved from: http://www.ers.usda.gov/media/134648/fts34501.pdf

Ministerio de Agricultura y Riego del Perú (2017). *Boletín Estadístico de Producción Agrícola, Pecuaria y Avícola Diciembre 2016.* Retrieved from: http://www.minagri.gob.pe/portal/download/pdf/herramientas/boletines/prod-agricola-pecuaria-avicola/2016/boletin-produccion-comercializacion-avicola-diciembre2016_020317.pdf

Oerlemans, L., Rooks, G. and Pretorius, T. (2005), Does technology and innovation management improve market position? Empirical evidence from innovating firms in South Africa. *Knowledge, Technology and Policy,* 18(3), 38-55.

Pellissier, R. and Nenzhelele, T. (2013). Towards a universal definition of competitive intelligence. *SA Journal of Information Management,* 15(2), Art.#559, 7 pages.

Porter, M. (1980). *Competitive strategy*. New York: Free Press.

Prescott, J. E. (1999). The evolution of competitive intelligence: designing a process for action. *Proposal management, APMP*, Vol. Spring, 37-52.

Prescott, J. and Miller, S. H. (2001). *Proven strategies in competitive intelligence.* New York: John Wiley & Sons, Inc.

Priporas, C., Gatsoris, L. and Zacharis, V. (2005). Competitive intelligence activity: evidence from Greece. *Marketing Intelligence & Planning,* 23(7), 659-669.

Rodrigues, V. Z. and Thome e Castro, L. (2017). Competitive intelligence within the telecommunication industry. Business practices in Brazil under the Wright-Pickton framework. *REGE - Revista de Gestao,* 24, 110-121.

Rodriguez Salvador, M. and Salinas Casanova, L. F. (2012). Applying Competitive Intelligence: the case of thermoplastics elastomers. *Journal of Intelligence Studies in Business,* 2(3), 41-47.

Ross, P., McGowan, C. and Styger, L. (2012). *A comparison of theory and practice in market intelligence gathering for Australian micro-businesses and SME's*. Proceedings of the 19th

International Business Research Conference, Monash University, Melbourne.

Rouach, D. and Santi, P. (2001). Competitive intelligence adds value: Five intelligence attitudes. *European Management Journal*, 19 (5), 552-559.

Salazar, K. O., Cardona, M. J. M., Ocampo, O. L. L. and Ovalle, A. M. C. (2014). Technology watch cycle analysis in the textile companies of south central Caldas. *Sciencia et Technica*, 19(1), 35-41.

Salles, M. (2006). Decision making in SMEs and information requirements for competitive intelligence. *Production Planning and Control,* 17(3), 229-237.

Smith, J, Wright, S., and Pickton, D. (2010). Competitive Intelligence programmes for SMEs in France: Evidence of changing attitudes. *Journal of Strategic Marketing,* 18(7), 523-536.

Smith, J. (2012). *Competitive intelligence behaviour and attitude antecedents in French small and medium sized enterprises in a funded intervention environment*. (Doctorate thesis). De Montford University, UK.

Soilen, K. S. (2013). An overview of articles on Competitive Intelligence in JCIM and CIR. *Journal of Intelligence Studies in Business,* 3(1), 44-58.

Soilen, K. S. (2016). A research agenda for intelligence studies in business. *Journal of Intelligence Studies in Business,* 6(1), 21-36.

Superintendencia Nacional de Aduanas y de Administracion Tributaria del Perú (2017). *Exportaciones P.A. 0806.10.00.00.* Retrieved from: http://www.aduanet.gob.pe/

Superintendencia Nacional de Aduanas y de Administración Tributaria del Perú (2018). *Anuarios.* Retrieved from: http://www.sunat.gob.pe

Tanev, S. and Bailetti, T. (2008). Competitive intelligence information and innovation in small Canadian firms. *European Journal of Marketing,* 42(7/8), 786-803.

Tang Tong, M. M. (2015). El rol de la inteligencia de mercados en las empresas exportadoras e importadoras en el Perú. *Ingeniería Industrial*, 33, 71-97.

Tarraf, P. and Molz, R. (2006). Competitive intelligence at small enterprises. *SAM Advanced Management Journal,* 71(4), 24–34.

Toker A, Seraj M., Kuscu A., Yavuz R., Koch S. and Bisson C. (2016). Social media adoption: a process based approach. *Journal of Organizational Computing and Electronic Commerce,* 26(4), 344-363.

Top Publications (2018). *Peru: the Top 10,000 companies.* Retrieved from: http://www.toponlineapp.com

Tryfonas, T. and Thomas, P. (2006). *Intelligence on competitors and ethical challenges of business information operations.* Proceedings 5th European Conference, Information Warfare and Security, June 1-2, National Defense College, Helsinki.

Villaroelg, C., Comai, A., Karmelic-Pavlov, V., Fernandez, A.O., and Arriagada, C.V. (2015). Design and implementation of a technological surveillance and competitive intelligence unit. *Interciencia,* 40(11), 751-757.

World Bank Group (2017). *Gaining Momentum in Peruvian Agriculture: Opportunities to increase productivity and enhance competitiveness.* Retrieved from: http://documents.worldbank.org/curated/en/107451498513689693/pdf/P162084-06-26-2017-1498513685623.pdf

World Bank (2017). *Gross Domestic Product.* Retrieved from: http://databank.worldbank.org

Wright, S. (2011). *A critical evaluation of competitive intelligence and insight management practice.* (Doctorate thesis). De Montfort University, UK.

Wright, S., Pickton, D. W. and Callow, J. (2002). Competitive Intelligence in UK firms: a typology. *Marketing Intelligence & Planning,* 20(6), 349-360.

Wright, S., Bisson, C. and Duffy, A. (2012). A Behavioural and Operational Typology of Competitive Intelligence Practices in Turkish SMEs. *Journal of Strategic Marketing,* 20(1), 19-33.

Wright, S., Bisson, C. and Duffy, A. (2013). Competitive Intelligence and Information Technology Adoption of SMEs in Turkey: Diagnosing Current Performance and Identifying Barriers. *Journal of Intelligence Studies of in Business*, 3(2), 5-29.

Zhan, X. and Chen, M. (2009). Competitive Intelligence monitoring in the risk prevention of SME's. *Journal of Service Science and Management,* 2(3), 230-235.

7. ACKNOWLEDGEMENT

The authors thank all participants who helped in the study and the Universidad de Lima, Peru which supported the fieldwork of this research.

8. APPENDİCES

Attitude	
Immune Attitude A1	Too busy thinking about today to worry about tomorrow. Thinks that the firm is either so small, so big or so special that it enjoys immunity from competitors and thus CI is a waste of time. Minimal or no support from either top management or other departments.
Task Driven Attitude A2	Finding answers to specific questions and extending what the firm knows about its competitors, usually on an ad-hoc basis. Departments more excited about CI than top management who don't see the benefits.
Operational Attitude A3	A process, revolving around the company as its centre, trying to understand, analyse and interpret markets. Top management usually trying to develop a positive attitude towards CI because they can see it might increase profit, and therefore personal bonuses. Unwilling or unable to think about the application of CI for the long term.
Strategic Attitude A4	An integrated procedure, in which competitors are determined as those who are satisfying our customer's needs, current and/or future. Monitoring their moves, anticipating what they will do next and working out response strategies. Receives both top management support, co-operation from other departments and is recognised by all as essential for future success.

Gathering

Easy Gathering G1	Firms which use general publications and/or specific industry periodicals and think these constitute exhaustive information. Unlikely to commit resources to obtain information which may be difficult or costly to obtain. Always looking for an immediate return on investment.
Hunter Gathering G2	Firms knowing that Easy Gathering information is available to all who care to look. Realise that if CI is to have a strategic impact then additional, sustained effort is required. Resources are available which allow researchers to access sources within reasonable cost parameters, back their instinct, follow apparently irrelevant leads, spend time talking, brainstorming and thinking about CI problems without always being pressured for "the answer". Firms which appreciate and support intellectual effort.

Use

Joneses User U1	Firms trying to obtain answers to disparate questions with no organisational learning taking place. Has commissioned a CI report from a consultant because that is what everybody else has done.
Knee Jerk User U2	Firms which obtain some CI data, fail to assess its quality or impact, yet act immediately. Can often lead to wasted and inappropriate effort, sometimes with damaging results. Such firms are most vulnerable to planted mis-information by competitors who are more CI aware.
Tactical User U3	CI used mostly to inform tactical measures such as price changes, promotional effort. Some firms can successfully argue that CI loses its impact and timeliness if it gets stuck at the strategic level but are, nevertheless, acutely aware of its potential value to the business.
Strategic User U4	CI is used to identify opportunities/threats in the industry and to aid effective strategic decision making. All levels of staff, both management and operational, are aware of CSF's and their attendant CI requirements. Continuous, legal measures are used to track competitors, simulate their strengths and weaknesses, build scenarios, and plan effective counter attacks. Decision makers are involved in a high number of "what-if?" discussions to which CI data is applied. Contingency planning and counter intelligence is a part of normal strategic thinking. Action plans are implemented and mistakes are seized upon as learning rather than firing opportunities. Open and facilitative management culture which displays trust and encourages involvement.

Location

Ad-Hoc Location L1	No dedicated CI unit. Intelligence activities, where undertaken are on an ad-hoc basis, subsumed into other departments, with intermittent or non-existent sharing policies.
Designated Location L2	Firms with a specific intelligence unit, full time staff, dedicated roles, addressing agreed strategic issues. Staff have easy access to decision makers, status is not a barrier to effective communication.

Technology Support

Simple Technology Support TS1	The company is just using the free web such as a search engine or looking at some web sites which require no specific knowledge. Also use general office software such as spreadsheet.
Average Technology Support TS2	Using “off the shelf” products such as meta-search engines which simply reorganise publicly available information for the firm use. The company might use web site which require specific knowledge (e.g. espacenet) and pay to use some specialised websites and databases (e.g. patent and finance).
Advanced Technology Support TS3	This information system holds vital and high level information as well as operational and tactical material. Is fully integrated across the business and continually evolves to meet the firm’s requirements. Content analysis (e.g. statistical analysis) provided.
High Technology Support TS4	In addition to advanced tools, firms use “clever” algorithms aimed at understanding automatically the competitive information collected. These algorithms are based on semantics.

IT Systems	
Dismissive IT Systems ITS1	Does not use any IT system to manage competitive information
Sceptic IT Systems ITS2	Has a system to manage competitive information but prefers to use paper based records. Does not trust IT systems sufficiently and is wary of their reliability
Standardized IT Systems ITS3	A standard existing system is purchased from a software vendor and installed on computers located within an organization.
Hosted IT Systems ITS4	A standard system is used, but it is not managed by the company itself (e.g. pay per view system).
Tailored IT Systems ITS5	In a tailored development, an off-the-shelf system or hosted solution is tailored according to an organization's needs regarding its competitive information.
Bespoke IT Systems ITS6	Unique to the firm system which has been designed in-house and aiming at collecting, analyzing and disseminating competitive information.

	Average FOB	Gathering			Attitude					Technology					Information System							Use					Location		
Cluster	$/ton**	G1	G2	Verdict	A1	A2	A3	A4	Verdict	TS1	TS2	TS3	TS4	Verdict	IS1	IS2	IS3	IS4	IS5	IS6	Verdict	U1	U2	U3	U4	Verdict	L1	L2	Verdict
Company size*																													
Big	2,336	3	5	G2	3	2	2	1	A1	7	0	1	0	TS1	4	2	1	0	0	1	IS1	3	0	1	4	U4	7	1	L1
Medium	2,518	6	9	G2	5	6	3	1	A2	11	1	3	0	TS1	6	2	4	2	0	1	IS1	8	0	0	7	U1	12	3	L1
Small & micro	2,259	2	12	G2	1	6	7	0	A3	13	1	0	0	TS1	8	1	1	1	0	3	IS1	2	0	0	12	U4	9	5	L1
% of grapes exports in total exports**																													
Higher than 75%	2,677	3	14	G2	3	8	6	0	A2	16	0	1	0	TS1	9	2	3	2	0	1	IS1	5	0	1	11	U4	12	5	L1
Lower than 75%	2,133	8	12	G2	6	6	6	2	A3	15	2	3	0	TS1	9	3	3	1	0	4	IS1	8	0	0	12	U4	16	4	L1
% of red globe in total grapes**																													
Lower than 50%	2,605	8	16	G2	6	9	7	2	A2	18	2	4	0	TS1	12	5	3	1	0	3	IS1	8	0	1	15	U4	19	5	L1
Higher than 50%	1,881	3	10	G2	3	5	5	0	A2	13	0	0	0	TS1	6	0	3	2	0	2	IS1	5	0	0	8	U4	9	4	L1
TOTAL	2,391	11	26	G2	9	14	12	2	A2	31	2	4	0	TS1	18	5	6	3	0	5	IS1	13	0	1	23	U4	28	9	L1
Grapes farm size																													
Lower than 100 hectares	1,932	2	3	G2	1	0	4	0	A3	5	0	0	0	TS1	2	1	1	0	0	1	IS1	2	0	0	3	U4	3	2	L1
between 101 and 500 hectares	2,413	7	13	G2	4	12	3	1	A2	18	1	1	0	TS1	11	2	3	2	0	2	IS1	7	0	0	13	U4	16	4	L1
Higher than 501 hectares	2,444	1	6	G2	2	1	3	1	A3	4	0	3	0	TS1	3	1	2	0	0	1	IS1	3	0	1	3	U1	6	1	L1
TOTAL	2,419	10	22	G2	7	13	10	2	A2	27	1	4	0	TS1	16	4	6	2	0	4	IS1	12	0	1	19	U4	25	7	L1

Sources:
* - Peru: Top Publications (2018)
** - Sunat (2016-2017)

JISIB
Journal of Intelligence Studies in Business
Vol. 8, No. 2 (2018) pp. 62-76
Open Access: Freely available at: https://ojs.hh.se/

Stevan Dedijer
1911-2004

Business intelligence through patent filings: An analysis of IP management strategies of ICT companies

Shabib-Ahmed Shaikh[a,b] and Tarun Kumar Singhal[c*]

[a] *Symbiosis International University (SIU), Lavale, Mulshi Taluka, Pune, Maharashtra, India*
[b] *CSIR-URDIP, Pune, Maharashtra, India*
[c] *IT Management, Symbiosis Centre for Management Studies, NOIDA, UP, India*

Corresponding author (*): tarun.singhal@sitm.ac.in

Received 29 August 2017 Accepted 21 August 2018

ABSTRACT Business intelligence enables enterprises to make effective and good quality business decisions. In the knowledge economy, patents are seen as strategic assets for companies as they provide a competitive advantage and at the same time ensure the freedom to operate and form the basis for new alliances. Publication or disclosure of intellectual property (IP) strategy based on patent filings is rarely available in the public domain. Because of this, the only way to understand IP strategy is to look at patent filings, analyze them and, based on the trends, deduce strategy. This paper tries to uncover IP strategies of five US and Indian IT companies by analyzing their patent filings. Gathering business intelligence via means of patent analytics can be used to understand the strategies used by companies in advocating their patent portfolio and aligning their business needs with patenting activities. This study reveals that the Indian companies are far behind in protecting their IPs, although they are now on course correction and have started aggressively protecting their inventions. It is also observed that the rival companies in the study are not directly competing with each other in the same technological domain. Different patent filing strategies are used by firms to gain a competitive advantage. Companies make use of disclosure as strategy or try to cover many aspects of a technology in a single patent, thereby signaling their dominance in a technological area and at the same time as they add information.

KEYWORDS Business intelligence, competitive intelligence, intellectual property, IPR, IP strategy, patent analytics, software patents

1. INTRODUCTION

Business intelligence helps enterprise users make effective and high quality business decisions. It includes multiple applications, tools and technologies for information gathering, accessing, and analyzing involving all factors that affects a business (Rajan, 2009). Howard Dressner, an analyst at the Gartner Group, first coined the term business intelligence in the early 1990s. Business intelligence has become the art of sifting through large amounts of data, extracting pertinent information, and turning that information into knowledge upon which timely actions can be taken. All successful enterprises have made use of business intelligence for their business (Chaudhuri, 2011).

As per Ranjan (2009), business intelligence reveals:

- The position of the firm relative to its competitors

- Changes in customer behavior and spending patterns
- The capabilities of the firm
- Market conditions, future trends, demographic and economic information
- The social, regulatory, and political environment
- What the other firms in the market are doing

Business intelligence as a strategic framework is becoming increasingly important in strategic management and in supporting business strategies (Alnoukari and Hananao 2017). Alignment between business and business intelligence strategies can be a powerful enabler of business strategy, including new business models that bring about organizational transformation (Watson & Wixom, 2007). Business intelligence using tacit knowledge can lead to intellectual capital including patents (Sveiby, 1997 ; Herschel & Jones, 2005).

The IT industry has grown rapidly since the 1960s, starting in the USA and has slowly become global (Cameron et al., 2006). Information and communication technology (ICT) innovations are usually incremental, fast changing and having a short lifecycle (Shaikh & Londhe, 2016). Firms investing in this continuously evolving technology expect quick returns for their investments by means of some protection. Intellectual property rights (IPR) are the rights given to persons over the creations of their intellect. The framework of IPR offers a wide range of protections such as patents, trademarks, copyright, design registrations, trade secrets, anti-competitive practices in contractual licenses, protection of new plant varieties and data protection.

A patent offers the strongest protection within the framework of IPR. It is a form of intellectual property granted by the government in order to secure legal protection for inventions by means of exclusive rights for a limited period in exchange for the public disclosure of an invention. Patents are also important for trade and industry worldwide as they attract foreign investment and rapid technology transfer (OECD, 2004). Patents also promote innovation by disclosing an invention in the public domain (Moser, 2005; Walaski, 2004). Patenting decisions are seen as important strategic considerations since gaining maximum value from a patent depends on the individual firm's ability to enforce the patent (Arrow, 1962 ; Dornelles, 2016).

Patents are a major source of information and when properly processed and analyzed, can yield a wealth of information on competitors' activities, R&D trends, emerging fields, and collaborations. Taking into account the filing practices (for example, broad or specific applications, filing routes, and territorial protection sought) associated with specific companies or domains, the analysis of patent portfolios can give a reasonably accurate idea of the volume of the activity in specific research areas, reveal underlying trends, detect emerging or hidden information or deviations from expected patterns, and more. Patent analysis can also yield a wealth of information related to research activity, collaborations, location of research work, key inventors and licensing (Grandjean et.al., 2005).

Strategic IP management can be offensive or defensive resulting in the formulation and execution of strategies related to technological IP, including issues such as how to acquire, create, govern, exploit and extract value from patents. Patents can also be used to understand technology and competitor intelligence (Holgersson 2012; Krig & Sandra, 2017).

Patenting usually has strong business co-relations (Pargaonkar, 2016). In the present study, patent filing data of selected ICT companies are used as a source of information for competitive/business intelligence to highlight the intellectual property (IP) management strategies of ICT companies. Patent landscape and the accompanying IP competitive intelligence involves understanding and anticipating the competitive environment within which a company operates. More specifically, IP competitive intelligence highlights emerging IP risks, provides patent portfolio benchmarking, monitors competitor technology development efforts, and predicts commercialization of technology (Pargaonkar, 2016). The main objective of IP competitive intelligence is to create value for competitive advantage. IP competitive intelligence improves decision quality and enables IP strategies by defining the relative competitive position. IP strategy becomes important when firms differentiate themselves using technology. In such cases, IP competitive intelligence analysis plays an important role for defining, creating and sustaining a winning IP strategy. IP competitive intelligence enables value creation and strengthens multiple aspects of an effective IP strategy (Pargaonkar, 2016).

Considering the above, there is a need to understand the various motives of firms to patent.

2. LITERATURE REVIEW

Various studies have been carried out in the field of competitive intelligence, business intelligence, their advantages to business in taking timely decisions, as well as the use of patent data for carrying out business intelligence for competitive advantage. Hughes (2017) reports that due to the high volume and speed of scientific research, it is impossible to collect, update and analyze the variables that impact the evolution of technologies as disruptive innovations need knowledge from adjacent technologies as well. Hughes (2017) proposes a model featuring expanded search depth, breadth and speed along with inputs from internal and external experts for identifying emerging technologies by coupling big data analytics machine learning with technology sequence analysis. On the other hand, Gauzelin and Bentz (2017) report on how small and medium-sized enterprises (SMEs) perceive and make use of business intelligence in decision making and highlight that business intelligence systems are perceived as a solution to various unforeseen disruptive events that hit the businesses unexpectedly. They report that assessing the success of business intelligence is not easy as they cover the entire organizations and their benefits are long term. SMEs lack business intelligence implementation due to a lack of financial and expertise capacity to implement it. However, small businesses deal with increasing volumes of data, hence making the appropriate choice of the best business intelligence in line with their strategy will allow them to have a competitive advantage. Collecting and analyzing data on business intelligence from SMEs, Gauzelin and Bentz (2017) report that business intelligence and its use have a far-reaching impact on the operation of SMEs. Søilen (2017), highlights the importance of competitive intelligence and market intelligence through a case study of two Swedish MNCs and reports that companies would succeed only if the competitive intelligence model, along with the specialist's role, are properly defined in bringing out and reporting facts instead of pleasing their seniors. Søilen (2017) also highlights that the expectations from the analysts is predicting the future, which at times is difficult. The analysts often also end up performing different tasks aside from analysis. With the increase in data and its low cost, competitive intelligence is largely defined by how well companies can draw conclusions from it, as the outcome is mainly dependent on the quality of data available and, at times of crisis, the demand for intelligence is the greatest.

Business intelligence can be viewed as a broader tool that includes knowledge management, enterprise resource planning, decision support systems and data mining (Gangadharan and Swamy, 2004). Business intelligence is also referred to as competitive intelligence, market Intelligence, customer intelligence, competitor intelligence, strategic intelligence or technical intelligence (Lönnqvist and Pirttimäki, 2006; Deshpande et.al, 2016). Scholars have define business intelligence as the process of collecting large amounts of heterogeneous data from multiple sources, analyzing that data using advanced analytical tools and methods, and quickly presenting a high-level set of reports to multiple users that condense the essence of that data into the basis of business actions, enabling management to make efficient and effective strategic business decisions that can help organizations to survive and thrive in the global economy (Stackowiak et al., 2007; Zeng et al., 2006; Ranjan, 2009).

The main challenge in any business intelligence solution is in its intelligence ability (Alnoukari and Hananao, 2017).

Business intelligence or competitive intelligence is considered to be an interdisciplinary field (Walker, 1994). Studies have suggested that competitive intelligence is associated with strategic management as well as knowledge management (Gabriel and Adiele, 2012; Calof and Viviers 2001) and intelligence has evolved as a discipline over time (Hoppe, 2015). Knowledge management can be perceived as an integral component of business intelligence (Herschel & Jones, 2005). It is usually defined in reference to collaboration, content management, organizational behavioral science, and technologies. Knowledge management is a systematic process of finding, selecting, organizing, distilling and presenting information in a way that improves an employee's comprehension in a specific area of interest (Herschel & Jones, 2005). It can be seen as consistent with resource-based theories of the firm, such as building and competing in a capability that could be quite difficult for others to imitate practically. Knowledge management was seen to be central to product and process innovation and improvement, to executive decision-making, and to organizational adaptation and renewal (Earl, 2001). Specific knowledge management activities help focus the organization on acquiring, storing and utilizing knowledge for such things as problem solving, dynamic learning, strategic planning and decision making. Alnoukari and Hananao (2017) report that the integration of business intelligence and corporate strategic management has a direct impact on modern and flexible organizations, which leads to a gain of competitive advantages as well as easier adatation to changing scenarios and corporate strategies.

The core advantage of any competitive intelligence system is to extract the knowledge needed about competitors' opportunities and threats (Alnoukari and Hananao, 2017). Competitive intelligence ensures a firm's competitiveness in the marketplace through a greater understanding of competitors and the overall competitive environment (Solomon, 2004). Competitive intelligence and market intelligence can also be built on competitors and influencers from exhibits and tradeshows (Solberg-Søilen, 2010).

Intellectual property assets are becoming increasingly important drivers of competitive advantage. This has forced organizations to effectively and efficiently mine their IP for business intelligence. Studies suggest that patent data is also a valuable source of competitive intelligence from which to derive a strategic advantage (Rouach and Santi, 2001; Dou et al., 2005; Grandjean et al., 2005; Shih et al., 2010; Deshpande et.al, 2016). Stern (2005) highlights that for creating competitive advantage, management must focus on exploiting IP during a product's lifecycle, which would encompass resource management and IP strategy. IP protection is a strategy that helps in formulating new strategies for protection of innovations and sustainable development. Patent data, its legal status and litigation data can be used for business intelligence purposes such as IP portfolio valuation, patent valuation, identification of competitors and their R&D efforts, assessment of active researchers in a particular field, assessment of patent quality, research quality, market trends, discover human capital, and to anticipate product launches (Sagacious Research, 2017). Patent analysis enables firms to make more informed decisions about their IP strategy and create value for their business (Great Dome Associates, 2018). Analysis of patent data accelerates innovation, saving time and money (Cubicibuc, 2017). A patent portfolio can be analyzed by carrying out patent landscaping (Tekic, 2014). Intellectual property landscaping is a strategic tool providing valuable business intelligence to ensure maximum understanding of the potential opportunities and competitive threats (hee.org, 2018). Patent landscaping provides insights which guide business strategies that include cost optimization, enforcement, licensing, R&D and mergers and acquisitions. Patent landscaping supports business strategies that help in the development of a quality patent portfolio, which in turn generates revenue and mitigates risk (ip.com, 2017).

IP strategy as a subset of the business strategy requires analysis of a firm's own inventive capabilities along with the IP landscape (Barrett,2005). A patent landscape can give a new perspective on a market by illustrating the players, their technologies and their filing history and behaviors over time. A comprehensive landscape informs companies about the strength of their IP and how it compares to other companies operating in the same market. Looking at IP in a broad perspective and applying business intelligence provides decision makers with actionable insights and a clear view of potential outcomes for various strategies (clearviewip, 2017).

Business intelligence is a systematic way of gathering data, analyzing and utilizing the same while making decisions in expanding, launching a new product, while carrying out mergers and acquisitions or for implementation of corporate strategies. Business intelligence from intellectual property rights helps organizations to follow a proactive approach (Siddhast.com). It provides information that will allow organizations to predict the behavior of their competitors, suppliers, customers, technologies, acquisitions, markets, products and services, and the general business environment with a degree of certainty (Vedder et al., 1999; Jourdan et al., 2008)

Stern (2005) reports that managing IP as a strategic driver helps businesses become market leaders, align their business strategy with product IP strategy and protect their technology via means of maintaining a product monopoly. This provides a competitive advantage, thereby encouraging and defining measures for IP evolution and exploitation. Wang (2011) highlights how patent intelligence can be used to make an intellectual property strategy. Citing various researchers, Wang (2011) reports that patent data can be used in core areas of technology management. Jürgensa and Solanab (2016) provide insights on the use of patent information for technology watch activities, classifying patent indicators for performance, technology, patent value and collaboration indicators. They report that to gain insights and competitive advantage in a specific technical domain, patent intelligence is used, which is also referred to as technology watch, technology intelligence or technology monitoring. This is a subdomain of competitive intelligence, a methodology for gathering analyzing and managing external information that can affect the organizations plans, decisions and operations. Citing various researchers Jürgensa and Solanab (2016), report that competitive intelligence through patent data allows one to measure current technical competitiveness and forecast technological trends in specific sectors. Highlighting a case study of the nanotechnology industry in Spain, Jürgensa and Solanab (2016) report that statistical analysis of patent information and its visualization is a powerful and successful way to gain insights into a technology that can be further used to monitor and evaluate technology activities.

Patents encourage and promote innovation by the disclosure of a technology in the public domain (Moser, 2005; Walaski, 2004). Patents also promote technology transfers and cross licensing. It is reported that countries that support stronger patent protection laws are much preferred destinations for foreign investments, new innovations and technology advancements (Goswami & Yadav, 2010; McGowan et al., 2007). Patenting does not always lead to a monopoly in pricing as it helps recover the R&D investment cost (Spinello, 2007) and hence the IP law allows the developer to profit from their creation (McGowan, Stephens & Gruber 2007). Increased incentives for patents have pushed firms towards "patent thickets" (Cockburn and MacGarvie, 2011). Patent thickets constitute a potentially imposing obstacle and do not allow freedom to operate for other businesses (Clarkson & Dekorte, 2006). Patent flooding and thickets have been used as anticompetitive tools to lock out competitors, especially in fast moving technological markets (Weatherall et al., 2013). The higher number of patent applications by firms also increases transactional costs and thereby opens the doors for strategic collaboration for patent pooling and cross-licensing so that the negative effects of patent thickets can be reduced (Zekos, 2006; Cockburn & MacGarvie, 2011). Patent laws have been interpreted over time to provide protection to the desired licensee. Even unwilling infringements by means of ignorance are not an excuse to avoid prosecution (Biles & Mann, 1992). Patent trolls have made an impact on business and innovation in the ICT sector. Trolls are becoming professional patent exploiters that have high quality technological patents (Pohlmann & Opitz, 2013). The trolls' blackmailing tactics can have adverse effects on the whole industry, which in turn may slow down innovation processes (Pohlmann & Opitz, 2013). Bessen & Hunt (2007) have warned that strategic patenting by non-R&D firms may pressurize firms to engage in a patent "arms race." However, Useche (2015) reports that a high number of patents reduces the risk of failure and acquisition, while quality increases their attractiveness as an acquisition target. Patents may give a firm an upper hand and a competitively advantageous position, thereby adversely affecting the competitor firms' market values (Chung et. al, 2016).

Large companies see IPRs as incentives to compete in IPR portfolios and patents as strategic assets to protect from competition, give design freedom, offer complementary protection and form a basis for new alliances. At the same time, SMEs see IPRs as restrictions and market barriers and they need to build their own IPR portfolio to make themselves more credible players in the market (Välimäki, 2001). One strategy followed by successful Chinese multinationals was to skip filling in the domestic market and go directly to developed countries by collaborating with the world's major companies, pointing out that high application does not result in profit (Nakai & Tanaka, 2010). Companies strongly involved in collaborating with customers that are experienced using patents are more inclined to use patents (Blind, 2007).

Among the many strategies used by companies, technology disclosures can be a rational offensive strategy to make its presence felt in a particular technological domain (Baker & Mezzetti 2005). This helps to make the patent office aware of its availability of potential prior art. This is done intentionally to create prior art that might stop rivals from patenting and making it more difficult to patent, hence extending the patent race through disclosure. Disclosing the intermediate results in a multi-stage patent context signals a firms' commitment to a research project, which may induce the rival to exit the competition or provide its followers ground to work ahead on the technology, depending on the knowledge spill over (Gill 2008). This at times leads to future acquisition or collaboration with its followers and at the same time prevents its competitors from working in the same domain.

Open source software (OSS) is attracting increasing commercial interest among firms as they take royalties over patented technologies of products and services sold as top-ups for OSS products (Fosfuri et al., 2008; Wen et. al., 2015). Firms with software patents highjack an OSS project and direct its development in a particularly favorable direction by threatening or exercising enforcement rights. Fosfuri et al. (2008) also states that patenting by firms that support OSS can also be for defensive purposes, thereby supporting their defensive strategies. Firms with large stocks of software patents or with large stocks of hardware trademarks are more likely to release OSS products (Fosfuri et al., 2008). Red Hat is making a profit from the sales, service and support of Linux even though Linux is open source (McGowan et al., 2007). It is seen that Red Hat has patent filings to protect its commercial interests (Shaikh & Londhe, 2016).

Firms patent not only to prevent imitation, but also to obtain bargaining power and improve their corporate image, to freely operate in the market, to extract value of their patents through licensing and royalties, to collaborate with technology leaders and to seek a competitive advantage. To strengthen a firm's technological leadership and to protect its innovation, patents serve as influential instruments of corporate strategy and have become an important source of competitive advantage (Grindley & Teece, 1997; Sullivan, 2001; Holgersson, 2012). Studies have pointed out the need for integrating and aligning patent strategy with a firm's business and technology strategy to generate valuable returns (Alexy et al., 2009; Granstrand, 2000; Smith & Hansen, 2002; Reitzig, 2004; Davoudi et.al., 2018; Lynskey 2009; Holgersson & Grandstrand, 2017).

The software market was born in the US and it still acts as a trendsetter for software patenting by opening its doors to software and business method patents (Cameron et al., 2006). Other countries are following the US to protect the interest of their researchers, as the failure to protect might affect a company's ability to operate freely at the basic level in the global market (Clarkson & Dekorte, 2006), which in turn would threaten their own existence (Dedrick & Kraemer, 1993; Jyoti et al., 2010). The best way to survive is to study and learn from the patenting strategies followed by the market leaders who are successfully protecting their inventions via means of patenting. Since no publication or public disclosure about IP strategies is available, the only way to understand such IP strategies is to look at the patent filings, analyze them and based on the trends, deduce their strategy. These insights thus obtained may help the IT industry to customize its strategy with respect to patent acquisition.

3. METHODOLOGY

The study covers patent data published from 2005 to 2014 from five Indian and five US ICT companies. The list of these companies is given in Table 1. The Derwent Innovation Database (https://clarivate.com/products/derwent-innovation/) was used to retrieve the relevant patent data for the study. The text mining and visualization tool Vantage Point (www.thevantagepoint.com) was used to clean, normalize and analyze the patent data. As the data retrieved was huge, it was also imported into a relational database for further filtering.

The search strategy consisted of assignee names of the ten firms. As the study was to find the technological trends and strategies, the patents searched were based on the application year (Trippe, 2015). The exemplary search strategy was:

CMP=("company names") AND
(AD>=(20050101) AND AD<=(20143112))

As patents are territorial in nature, the same invention may be duplicated by way of multiple fillings in different countries, which can be referred to as patent families. To reduce this form of duplication, one representative of each family was retained to obtain the data set highlighted in Table 1.

The bibliographic details of patents such as the title, abstract, claim, priority date, assignee name, inventor name, INPADOC family members, and citations have been used for the analysis.

Table 1 Companies with patent data sets and patent families.

Company	Patent Data Set	Patent Families
International Business Machines Corp.	87,086	24,206
Samsung Electronics Co. Ltd.	168,170	26,885
Microsoft Corp.	118,860	19,274
Google Inc.	57,589	8,931
Qualcomm Inc.	179,640	13,899
Tata Consultancy Services Ltd	1,803	414
Infosys Ltd	644	273
Wipro Ltd	799	375
HCL Technologies Ltd	316	205
Mahindra IT & Business Services	523	263

4. ANALYSIS AND VISUALIZATION

4.1 Patenting trends for US and Indian IT companies

The overall patenting activity for these US and Indian IT companies between 2005 and 2014 can be seen in Figure 1 and Figure 2, respectively. The figures highlight that the patenting activity of the US companies is higher than their Indian counterparts, which lag in protection of software innovations. The US companies applied for about 93,000 patents, while the Indian companies applied for less than 2% of that quantity, with about 1500 patent applications in the same time period.

It is observed that the patent applications of Google and Qualcomm have gradually increased in the study period, while that of Microsoft decreased. Samsung leads the application rate for almost 5 years, with more than 3,000 patents each year. On the other hand, Indian companies such as TCS, HCL and Wipro aggressively started patenting their activities only in 2010, 2011 and 2012, respectively. Infosys and Mahindra made their

Figure 1 : Patent Application Trend of US companies

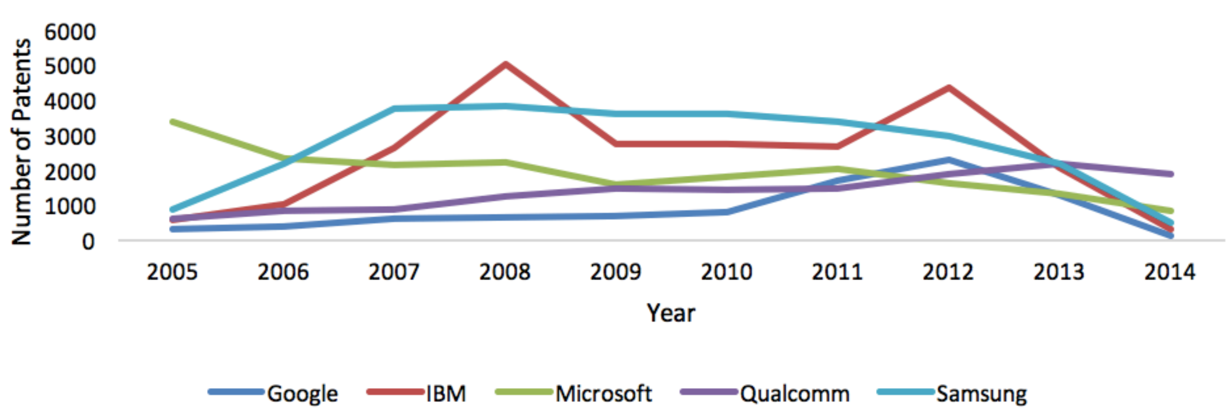
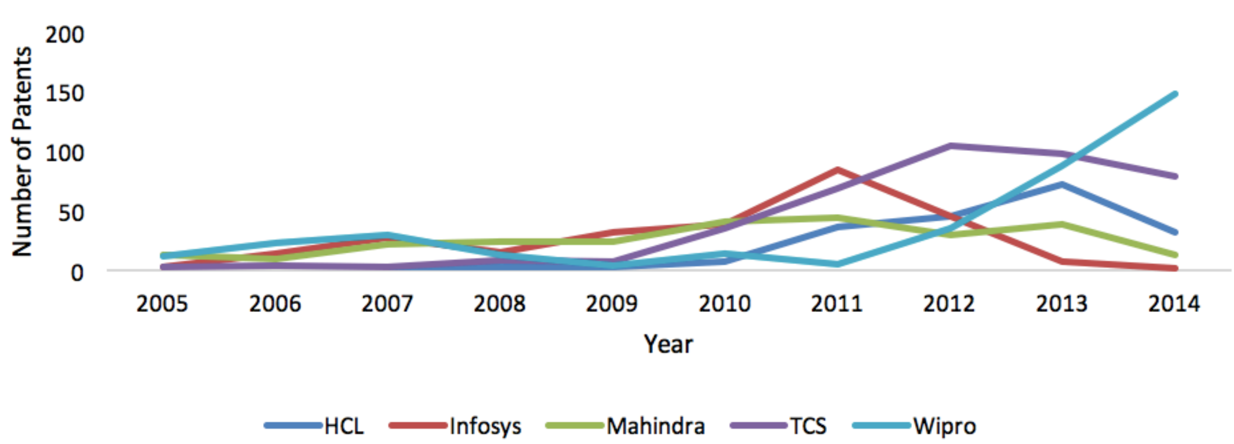
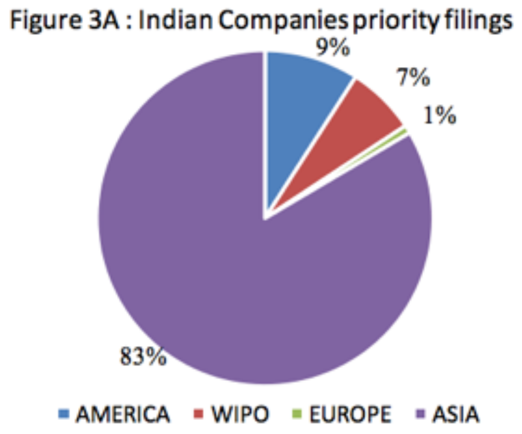
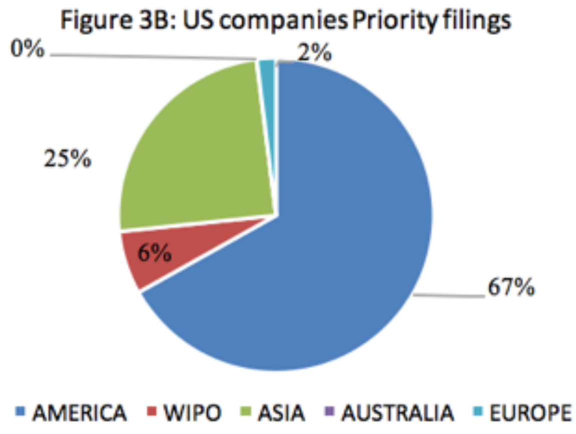


Figure 2 : Patent Application Trend of Indian companies







presence felt throughout the decade under consideration. After comparing the Indian and the US firms it can be said that the Indian companies entered late into the patenting foray.

4.2 Origin of Inventions for US and Indian IT companies

The origin of an invention can be found by using patent data (Trippe, 2015). The priority filing country in the patent document is considered to be an indicator for the origin of a particular invention, as companies usually prefer to first file for a patent in the same country in which the technology is invented. Figure 3 illustrates the priority country filing trends for the Indian and US IT companies.

During the study, it was observed that the majority of the patents (67%) claim the US as the priority country. However, a closer look revealed that the Indian companies have India as their origin of invention. A further analysis of the top filers from India reveals Wipro has its patent origins in at least 9 countries while TCS has its origin of invention in 4 countries, Infosys and HCL in 3, and Mahindra had its origin of inventions in 2 countries. The study of major US filers reveals that Samsung leads the way by priority filing around 78% of its patents first in Korea followed by 18% in the US. Samsung and Qualcomm have priority filings in at least 12 countries and 78% of Qualcomm's, and 86% of Microsoft's, inventions originated from the US. Microsoft has filings for origin of inventions from 13 countries. IBM has a spread across 14 countries and has about 88% of its inventions' priority filings in the USA. Around 5% of IBM's inventions originate in Europe. The reason that the US-based companies have many countries as their origins of invention can be attributed to their global presence in the form of technology and R&D centers in multiple countries, along with their collaboration in research. However, this is not the case of the Indian companies, as they operate in selected markets other than India such as the USA and Europe only. Wipro is the only Indian company with priority filings for inventions from at least 9 countries. It is also interesting to note that WIPRO has around 16% of its patents in the USA and 3% of patents originating in Singapore.

4.3 Patent legal status for US and Indian IT companies

Patents' legal statuses are an important component of patent information. They show whether a patent is dead or alive. They can also throw light on the various strategies used by the patenting firms, such as which technology is still protected and where, or whether it will soon become freely available in the public domain (WIPO-a). Alive patents are the ones that are valid and can be enforced. The dead patents are the ones whose applications are either withdrawn, rejected or the granted patent has expired, lapsed or been revoked for various reasons such as non-payment of maintenance fees. There is also a third category in the legal status known as "indeterminate," where patents are assumed to be applications undergoing examination, the examination is pending or whose status is not known.

Table 2 highlights the legal status of patents in percentage for the 10 companies studied. It is interesting to note that Infosys has around 92% of its patents live and enforceable. Inversely, about 30% of IBM's patents are unenforceable due to withdrawal of the application, rejection, lapse or revocation. This may be seen as an offensive tactic by IBM to make data public via means of disclosure to force firms out of competition and at the same time save costs incurred on prosecution or maintenance of patents. It might also be due to the technology in IBM's patents becoming absolute. Even then a figure of 30% is quite high.

Table 2 Patent legality status in percentage.

Company Name	Living Patents	Dead Patents	Indt.
IBM	67	30	3
Samsung	85	9	6
Microsoft	81	9	10
Qualcomm	74	10	16
Google	82	6	12
TCS	77	3	20
HCL	9	0	91
Mahindra	26	0	74
Infosys	92	2	6
Wipro	41	0	59

4.4 Technological trends of US and Indian IT companies.

The International Patent Classification (IPC) is used in a patent document to classify the patent according to the technical fields it claims. An analysis into the top IPC-4 digit for the 10 companies studied revealed that 7 companies (IBM, Microsoft, Google, HCL, Infosys, TCS and Wipro) lead with maximum patenting in G06F which indicates "Electrical Digital Data Processing" (Table 3).

Around 62% of Microsoft's patents were in the IPC-4 digit class G06F, while IBM has around 55% of its patents in G06F and about half of Google's patents were in G06F IPC-4 digits. Samsung lead with the majority of their patents in H01L, with 17% of its total filings in the class indicating "Semiconductor Devices" while Qualcomm has about 34% of its patent filings in H04W, indicating "wireless communication networks" and Mahindra with 9% in B60R, or "Vehicles".

Table 3 Count of Patents for Top IPC-4 digits of each companies.

	IPC-4 Digit			
	G06F	H01L	H04W	B60R
Google	4456			
IBM	13192			
Samsung		4519		
Microsoft	11889			
Qualcomm			4750	
HCL	91			
TCS	184			
Infosys	187			
Wipro	188			
Mahindra				24

Table 4 Count of patents for the top 3 IPCs of each company.

IPC	Google	IBM	Samsung	Microsoft	Qualcomm	HCL	TCS	Infosys	Wipro	Mahindra
G06F001730	1510	2138	-	2900	-	9	61	37	-	-
G06F001516	685	1544	-	1675	-	-	-	27	-	-
G06F000700	481	-	-	-	-	-	-	-	-	-
G06F000944	-	1540	-	1548	-	-	-	43	-	-
H04W000400	-	-	924	-	912	-	-	-	-	-
H01L002100	-	-	504	-	-	-	-	-	-	-
G09G000500	-	-	437	-	-	-	-	-	-	-
H04L002906	-	-	-	-	520	-	-	-	-	-
H04L000100	-	-	-	-	508	-	-	-	-	-
G06Q001000	-	-	-	-	-	9	-	-	-	-
G06F000944	-	-	-	-	-	7	27	-	-	-
G06F001700	-	-	-	-	-	-	-	-	14	-
G06Q001000	-	-	-	-	-	-	-	-	13	-
G06F001730	-	-	-	-	-	-	-	-	12	-
G06Q001006	-	-	-	-	-	-	29	-	-	-
B60R002100	-	-	-	-	-	-	-	-	-	4
B05B001500	-	-	-	-	-	-	-	-	-	3
B60K002000	-	-	-	-	-	-	-	-	-	3

Table 5 Claim count in patent applications.

Claim Count	Google	IBM	Samsung	Microsoft	Qualcomm	HCL	TCS	Infosys	Wipro	Mahindra
0-10	833	48	7341	6890	1347	200	169	31	308	258
11-20	4539	15349	13533	11448	2844	27	234	128	90	9
21-30	3026	4144	4599	689	3717	12	20	90	74	0
30-50	757	415	1289	214	4179	6	6	25	8	0
51-75	107	16	117	28	1374	0	2	1	1	0
76-100	27	4	11	4	336	0	0	0	0	0
>100	11	0	2	1	103	0	0	0	0	0

Table 6 Illustration of counts for family size, claim count, citations, number of inventors and assignee count. *Rounded off to the nearest whole digit.

	IBM	Samsung	Microsoft	Qualcomm	Google	TCS	HCL	Mahindra	Infosys	Wipro
Average Family Country*	4	6	6	13	6	4	1	2	2	2
Average Claim Count*	17	16	15	33	21	14	19	6	21	20
Maximum Claim Count	99	126	113	208	119	59	50	18	56	54
Minimum Claim Count	1	1	1	1	1	1	3	1	1	5
Maximum Assignee Count	62	23	19	21	16	13	6	5	11	9
Average Assignee Count*	3	2	2	2	2	2	1	1	3	1
Maximum Backward Reference	571	598	1248	1509	2007	37	15	26	148	51
Minimum Backward Reference	1	1	1	1	1	1	1	2	1	1
Average Backward Reference*	23	16	27	29	26	6	5	10	12	9
Maximum Forward Reference	112	151	189	85	206	32	6	50	101	28
Minimum Forward Reference	1	1	1	1	1	1	1	1	1	1
Average Forward Reference*	4	5	9	4	9	3	2	19	5	5
Maximum Inventor Count	61	21	60	20	29	12	10	8	10	8
Average Inventor Count*	4	3	4	3	3	3	3	3	3	2

A further analysis of the IPCs taking into consideration the full IPC revealed that all these companies are working in different domains with a minimum domain mapping with each other. This is highlighted in Table 4 for the top 3 patenting technologies of each company based on the IPC. Google has around 17% of its technologies patented in G06F001730 ("information retrieval"), while Microsoft and IBM map the same technology with around 15% and 9% of their total patents, respectively. Qualcomm and Samsung work in the same domain of H04W000400 ("services specially adapted for wireless communication networks") with about 7% and 3% of their total patent filings in this domain.

Even though Microsoft leads in G06F001730 ("information retrieval"), a closer look of its filings reveals that it has decreased its applications in information retrieval for the last 10 years. At the same time, Google has increased its activity in this field.

4.5 Claims filed by US and Indian IT companies

Claims play an important role in patent document. The patent description reveals how to make and use the invention, while the claims define the scope of legal protection and provide boundaries of the patent owner's exclusive rights. Hence, patent assertion for novelty depends on its claims (Merges & Nelson, 1990). Thus the number of claims of a patent document determines the depth and breadth of the technology for which protection is sought.

Table 5 shows the claim counts for each of the companies in this study. Usually, patent claims are in the range of 1-10, as claims above 10 incur additional filing charges. However, as seen above, less than 20% of the patents have claim counts of less than 10. Around 53% of the patents have claim counts between 11 and 20. All of the US-based companies have the maximum patent applications with claims in the range of 11-30. One important thing to note is that the US-based companies also have about 3% of their patents with more than 50 claims in a patent document. At the same time, Qualcomm has more than 13% of its patents with more than 50 claims each. Google has 11 patents with claim counts of over 100, Qualcomm has 103 (about 1%) patents with claim counts of over 100. This is much higher than the average patent claim counts. Google can be seen in Table 5 with a patent having 119 claims, whereas Qualcomm had a patent (application number EP2559309A1) with 208 claims. It can also be seen that Qualcomm leads with the highest average claim count in patents with more than 33 and Google following it with an average of about 21 claims per patent document. Indian companies Infosys, Wipro and HCL have an average of around 20 claims per patent document.

4.6 Analysis based on patent family size, claims count, number of citations, number of inventors and assignees for Indian and US IT companies

As highlighted in Table 6, the average family size of a patent is 4.7, while all the Indian companies are below this average count, US companies, baring IBM, have an average family size per patent higher than 6. Qualcomm has the highest average family, around 13 per patent. Base on this, it can be derived that Qualcomm tries to enforce its inventions in most countries simultaneously. However, IBM, which has much higher patent families than Qualcomm, has an average family size of around 4. This is the lowest for the US-based companies. If correlated with the origin of inventions, IBM has the maximum presence, in 13 countries, from where its technology has emerged. Hence it can be deduced that IBM's strategy is to enforce particular technologies in specific countries only and not in many countries, as in the case of Qualcomm.

A patent application contains references to other patent documents in its description (WIPO-b). These references can be forward or backward references. While the backward citations refer to the publicly available technological documents to form prior-art, the forward citations highlight all other patents and refer to the new patent application (WIPO-c). These citations, when analyzed, give insights into the evaluation of a particular technology (Breitzman, 2010).

Table 6 shows that all of the US based companies have an average backward citation above 20, except for Samsung which has an average citation above 16. With respect to the Indian companies, the average backward citation is less than 10. The US-based companies had at least one patent with a maximum backward citation of more than 500. Google had a patent with 2007 citations, whereas Qualcomm and Microsoft have patent publication with maximum backward citations of 1509 and 1248, respectively.

The forward citations are also useful from a competitive or business intelligence perspective to identify players working in a similar area or technology to the new patent application. Monitoring the forward citations of a new patent application allows a user to identify new competitors entering a similar field of technology, potential infringers and possibly, potential licensing opportunities (Minesoft). Google and Microsoft have the highest average forward citations for patents, with an average of about 9 forward references per patent, while HCL had the minimum with 2. Thus it can be inferred that patents of Google and Microsoft are used by other players to advance their technologies. Google has a patent with the maximum of 206 forward citations, while Microsoft has 189 forward cited patents for its publication. Infosys tops the list on the Indian side with 101 forward references in its patent publication number US7787887B2.

The number of inventors per patent is summarized in Table 6. It can be seen that for all of the companies the average inventor count per patent is around 3. Even then, IBM and Microsoft have patents with inventor counts of more than 60, and they are the only two companies with an average inventor count around 3.5.

5. CONCLUSION

Business intelligence in general and competitive intelligence in particular has been traditionally used for inputs related to sales, marketing and finance. However, the use of patents as strategic business tools has opened a new horizon for the use of patent analytics in gaining inputs based on business intelligence and competitive intelligence. Patent analytics based on competitive intelligence can be used for understanding the strategies used by companies in advocating their patent portfolio and aligning their business with patenting activities.

It can be seen from the study that the ICT companies in the study are not directly competing with each other in the same technological domain, except for G06F001730 (information retrieval). Indian companies are far behind in protecting their IP, although they are now on course correction and have started aggressively protecting their inventions. It is observed that the patent filing strategy of Qualcomm differs from its competitor IBM because Qualcomm is filing patents in all major countries while IBM has it presence felt only in specific countries, which can be seen from average patent family countries count. Claims in the patent document highlight the technological depth and breadth of patent applications, and Qualcomm seeks protection to maximum claims, thereby revealing its strategy of covering many aspects of a technology within a single patent application. Based on forward and backward citations, it appears that Microsoft and Google possess high quality patents. It is apparent that IBM uses disclosure strategies, as 30% of IBMs patents are dead, resulting in the technology coming into the public domain. This may be a tactic to force competitors out of their activities. Contrary to IBM's tactics, Samsung has 85% of its patents enforced, while rentaining the highest number of patent families, proving it to be a serious player in protecting its intellectual property. Business and competitive intelligence, when use to study IP competitive analysis, can yield IP strategies that may enable firms to align their IP strategy with their business strategy.

6. REFERENCES

Alnoukari, M. and Hananao, A. (2017) Integration of business intelligence with corporate strategic management. Journal of Intelligence Studies in Business. 7 (2) 5-16.

Baker, S., & Mezzetti, C. (2005). Disclosure as a Strategy in the Patent Race*. Journal of Law and Economics, 48(1), 173-194

Barrett, W. A. (2005). Building a strategy for maximizing intellectual property value. Nature biotechnology, 23(3), 387.

Bessen, J. & Hunt, R. M. (2007). An Empirical Look at Software Patents. *Journal of Economics & Management Strategy*, Volume 16, Number 1, 157–189.

Biles, G. E., & Mann, D. A., (1992). Software Law and Its Perils : Copyright, Patent, and the Racketeering Statute. *Information Systems Security*, 1:1, 42-49.

Björn Jürgensa* and Victor Herrero-Solanab; (2016); Patent bibliometrics and its use for technology watch ; Journal of Intelligence Studies in Business ; Vol. 7, No. 2 (2016) pp. 17-26

Blind, K. (2007). Intellectual Property in Software Development: Trends, Strategies and Problems. Review of Economic Research on Copyright Issues, 4(1), 15-25.

Breitzman, A. (2010). Why do Inventors Reference Papers and Patents in their Patent Applications?. *IEEE*. https://www.ieee.org/documents/ieee_why_inventors_reference.pdf

Cameron, D. M., MacKendrick, R. S., & Chumak, Y. (2006). Patents For Computer Implemented Inventions And Business Methods. *Canadian IT Law Association*; Toronto, 2006.

Calof, J.L. & Viviers, W. (2001). Adding competitive intelligence to South Africa's knowledge management mix. Africa Insight 31(2): 61-67.

Chaudhuri, S., Dayal, U., & Narasayya, V. (2011). An overview of business intelligence technology. Communications of the ACM, 54(8), 88-98.

Chung, S., Han, K., Animesh, A., & Pinsonneault, A. (2016, January). Competitive Impacts of Software Patents in the IT Industry. In *System Sciences (HICSS), 2016 49th Hawaii International Conference on* (pp. 5190-5199). IEEE.

Clarkson, G., & Dekorte, D., (2006). The Problem of Patent Thickets in Convergent Technologies. *New York Academy of Sciences,* 1093, 180–200.

ClearViewIP (2017), Where the Value Exists in a Patent Landscape ; accessed online, source : http://www.clearviewip.com/patent-landscape/

Cockburn, I., M., & MacGarvie, M. J., (2011) Entry and Patenting in the Software Industry. Management Science 57(5):915-933

Cubicibuc (2017); Strategic IP Management ; accessed online, source : http://www.cubicibuc.com/strategic-ip-management

Davoudi, S. M. M., Fartash, K., Zakirova, V. G., Belyalova, A. M., Kurbanov, R. A., Boiarchuk, A. V., & Sizova, Z. M. (2018). Testing the Mediating Role of Open Innovation on the Relationship between Intellectual Property Rights and Organizational Performance: A Case of Science and Technology Park. Eurasia Journal of Mathematics, Science and Technology Education, 14(4), 1359-1369.

Dedrick J. & Kraemer, K. L., (1993). Information Technology in India: The quest for self-reliance. *Asian Survey*, Vol. 33, No. 5 463-492.

Deshpande, N., Ahmed, S., & Khode, A. (2016). Business intelligence through patinformatics: A study of energy efficient data centres using patent data. *Journal of Intelligence Studies in Business*, *6*(3).

Dou, H., Leveillé, V., Manullang, S., & Dou Jr, J. M. (2005). Patent analysis for competitive technical intelligence and innovative thinking. *Data science journal*, *4*, 209-236.

Dornelles, J. (2016). Why are they hiding? Patent secrecy and patenting strategies.; Available at SSRN: https://ssrn.com/abstract=2802153 or http://dx.doi.org/10.2139/ssrn.2802153

Earl, M. (2001). Knowledge management strategies: Toward a taxonomy. Journal of management information systems, 18(1), 215-233.

Financial Express; 06 Jul 2013;Pg 7. ; INDIA BUSINESS INSIGHT

Fosfuri, A., Giarratana, M, S. & Luzzi, A. (2008). The Penguin has entered the building: The commercialization of open source software products. *Organization Science*, Vol. 19, No. 2, March–April 2008, 292–305.

Gabriel, J.M.O. & Adiele, K.C. (2012). Competitive Intelligence as panacea for environmental vagaries in Nigeria. Economic Journal of A 2 Z 1(1): 25-30.

Gangadharan, G. R., & Swami, S. N. (2004, June). Business intelligence systems: design and implementation strategies. In Information Technology Interfaces, 2004. 26th International Conference on (pp. 139-144). IEEE.

Gill, D. (2008). Strategic disclosure of intermediate research results. Journal of Economics & Management Strategy, 17(3), 733-758.

Goswami, G., & Yadav, K.P., (2010), Software Patents and the Current Trends, *International Journal of Computer Science & Communication*, Vol. 1, No. 1, 39-41.

Grandjean, N., Charpiot, B., Pena, C. A., & Peitsch, M. C. (2005). Competitive intelligence and patent analysis in drug discovery: Mining the competitive knowledge bases and patents. Drug Discovery Today: Technologies, 2(3), 211-215.

Great Dome Associates (2018), The heart of the innovation Economy is Intellectual Property ; accessed online, source : http://www.great-dome.com/intellectual-property-strategy/

HEE (2018), IP landscaping ; accessed online, source : https://www.hee.org.uk/services/explore/ip-landscaping

Herschel, R. T., Jones, N. E., (2005) "Knowledge management and business intelligence: the importance of integration", Journal of Knowledge Management, Vol. 9 Issue: 4, pp.45-55

Holgersson, M., & Granstrand, O. (2017). Patenting motives, technology strategies, and open innovation. Management Decision, 55(6), 1265-1284.

Hoppe, M. (2015) Intelligence as a discipline, not just a practise. *Journal of Intelligence Studies in Business*. Vol 5, No 3. Pages 47-56.

Hughes, S. F. (2017). A new model for identifying emerging technologies. *Journal of Intelligence Studies in Business*, *7*(1).

IP.Com (2017) ; How to Turn Patent Search Into Business Intelligence ; accessed online, source : http://ip.com/blog/turn-patent-search-business-intelligence/

Jourdan, Z., Rainer, R. K., & Marshall, T. E. (2008). Business intelligence: an analysis of the literature. *Information Systems Management*, *25*(2), 121-131.

Jyoti , Banwet, D.K., & Deshmukh, S.G. (2010). Modelling the success factors for national R&D organizations: a case of India. *Journal of Modelling in Management*, Vol. 5 No. 2, 2010, 158-175.

Krig, M. L., & Sandra, L. (2017). Business Value Enhancing Factors of Aligning IP Strategy with Corporate Strategy; available at : http://www.diva-portal.org/smash/get/diva2:1116900/FULLTEXT02

Lönnqvist, A., & Pirttimäki, V. (2006). The measurement of business intelligence. Information systems management, 23(1), 32

McGowan, M. K., Stephens, P., & Gruber, D., (2007). An exploration of the ideologies of software intellectual property: The impact on ethical decision making. *Journal of Business Ethics*, Vol. 73, No. 4, 409-424.

Merges, R. P. & Nelson, R.R., "On the complex economics of patent scope," Columbia Law Review, Vol 90, no. 4, pp.839-916, May 1990.

Minesoft, The Power Of Patent Citations ; Accessed from Minesoft website on March 2018; source : https://minesoft.com/2016/02/19/the-power-of-patent-citations/

Moser, P., (2005). How Do Patent Laws Influence Innovation? Evidence from Nineteenth-Century World's Fairs. *The American Economic Review*, Vol. 95, No. 4, 1214-1236.

Nakai, Y., & Tanaka, Y. (2010, July). Chinese company's IPR strategy: How Huawei Technologies succeeded in dominating overseas market by Sideward-Crawl Crab Strategy. In Technology Management for Global Economic Growth (PICMET), 2010 Proceedings of PICMET'10: (pp. 1-5). IEEE.

OECD (2004), Patents And Innovation: Trends And Policy Challenges - Organisation For Economic Co-Operation And Development.

Pargaonkar, Y. R. (2016). Leveraging patent landscape analysis and IP competitive intelligence for competitive advantage. *World Patent Information*, *45*, 10-20.

Pohlmann, T., & Opitz, M., (2013). Typology of the patent troll business, *R&D Management* 43, 2, 103-120.

Ranjan, J. (2009). Business intelligence: Concepts, components, techniques and benefits. Journal of Theoretical and Applied Information Technology, 9(1), 60-70.

Rouach, D., & Santi, P. (2001). Competitive Intelligence Adds Value: Five Intelligence Attitudes. *European management journal*, *19*(5), 552-559.

Sagacious Research (2017), accessed online, source : http://sagaciousresearch.com/blog/how-use-patent-data-business-intelligence/

Shaikh, S. A., & Londhe, B. R., (2016). Intricacies of Software Protection: A Techno-Legal Review; *Journal of Intellectual Property Rights*, Vol. 21; 157 - 165

Shih, M. J., Liu, D. R., & Hsu M. L. (2010). Discovering competitive intelligence by mining changes in patent trends. *Expert Systems with Applications*, *37*(4), 2882-2890

Siddhast.com ; IP Business Intelligence & Analytic ; accessed online, source : http://siddhast.com/service/ip-business-intelligence-analytic/

Søilen, K. S. (2010). Boosting innovation and knowledge through delocalization: market intelligence at trade shows. Problems and Perspectives in Management, 8(3), 200-207

Søilen K. S. (2017); Why care about competitive intelligence and market intelligence? The case of Ericsson and the Swedish Cellulose Company; *Journal of Intelligence Studies in Business;* Vol. 7, No. 2 (2017) pp. 27-39

Solomon, N., (2004) "Business Intelligence," Communications of the Association for Information Systems: Vol. 13 , Article 15

Sophian Gauzelin* and Hugo Bentz; (2017) ; An examination of the impact of business intelligence systems on organizational decision making and performance: The case of France; Journal of Intelligence Studies in Business; Vol. 7, No. 2 (2017) pp. 40-50

Spinello, R. A., (2007). Intellectual property rights. *Library Hi Tech*, Vol. 25 No. 1, 12-22.

Stackowiak, R., Rayman, J., & Greenwald, R. (2007). *Oracle data warehousing & business intelligence SO*. John Wiley & Sons.

Stern, A. (2005). Leveraging intellectual property for strategic advantage in product development. *South African Journal of Information Management*, *7*(4), 1-1.

Sveiby, K. E. (1997). *The new organizational wealth: Managing & measuring knowledge-based assets*. Berrett-Koehler Publishers.

Tekic, Z., Drazic, M., Kukolj, D., & Vitas, M. (2014). From patent data to business intelligence–PSALM case studies. *Procedia Engineering*, *69*, 296-303.

Trippe, A. (2015). Guidelines for preparing patent landscape reports. *Patent landscape reports. Geneva: WIPO*, 2015.

Useche, D., (2015). Patenting Behaviour and the Survival of Newly Listed European Software Firms, Industry and Innovation, 22:1, 37-58

Välimäki, M. (2001). Strategic Use of Intellectual Property Rights in Digital Economy–Case of Software Markets. Helsinki Institute for Information Technology, online http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.196.5665&rep=rep1&type=pdf

Vedder, R. G., Vanecek, M. T., Guynes, C. S., & Cappel, J. J. (1999). CEO and CIO perspectives on competitive intelligence. *Communications of the ACM*, *42*(8), 108-116.

Walaski, J. (2004). Software patents - are they a help or hindrance to innovation?. *IEE Engineering Management*, , 42-43.

Walker, T. D. (1994). The literature of competitive intelligence. Available at : https://www.ideals.illinois.edu/bitstream/handle/2142/7958/librarytrendsv43i2i_opt.pdf?sequence=1

Wang, X. (2011). Patent intelligence and business strategy. *African Journal of Business Management*, *5*(10), 3935-3941.

Watson, H. J., & Wixom, B. H. (2007). The current state of business intelligence. Computer, 40(9).

Weatherall, K & Webster, E., (2013), Patent Enforcement: A Review of the Literature; Journal of Economic Surveys (2013) Vol. 00, No. 0, pp. 1–32

Wen, W., Ceccagnoli, M., & Forman, C. (2015). Opening up intellectual property strategy: Implications for open source software entry by start-up firms. *Management Science*, *62*(9), 2668-2691.

WIPO-a ; Accessed from WIPO website on February 2018 source : http://www.wipo.int/patentscope/en/programs/legal_status/

WIPO-b ; Accessed from WIPO website on February 2018 source : http://www.wipo.int/export/sites/www/cws/en/taskforce/citation_practices/docs/epo_citation_practice_summary.pdf

WIPO-c ; Accessed from WIPO website on February 2018 ; Handbook on Industrial Property Information and Documentation; Ref.: Standards – St.14 Page: 3.14.1 Recommendation for the Inclusion of References Cited In Patent Documents; source : http://www.wipo.int/export/sites/www/standards/en/pdf/03-14-01.pdf

Zack Jourdan , R. Kelly Rainer & Thomas E. Marshall (2008) Business Intelligence: An Analysis of the Literature , Information Systems Management, 25:2, 121-131,

Zekos, G. I., (2006). Software Patenting. *The Journal of World Intellectual Property,* Vol. 9, no. 4, 426–444.

Zeng, L., Xu, L., Shi, Z., Wang, M., & Wu, W. (2006, October). Techniques, process, and enterprise solutions of business intelligence. In *Systems, Man and Cybernetics, 2006. SMC'06. IEEE International Conference on* (Vol. 6, pp. 4722-4726). IEEE.

JISIB

Journal of Intelligence Studies in Business
Vol. 8, No. 2 (2018) pp. 77-84
Open Access: Freely available at: https://ojs.hh.se/

Stevan Dedijer
1911-2004

Business intelligence for social media interaction in the travel industry in Indonesia

Michael Yulianto[a], Abba Suganda Girsang[a*] and Reinert Yosua Rumagit[a]

[a] *Computer Science Department, BINUS Graduate Program-Master of Computer Science, Bina Nusantara University, Jakarta, Indonesia 1148*

Corresponding author (*): mailto:agirsang@binus.edu

Received 14 May 2018 Accepted 25 July 2018

ABSTRACT Electronic ticket (eticket) provider services are growing fast in Indonesia, making the competition between companies increasingly intense. Moreover, most of them have the same service or feature for serving their customers. To get back the feedback of their customers, many companies use social media (Facebook and Twitter) for marketing activity or communicating directly with their customers. The development of current technology allows the company to take data from social media. Thus, many companies take social media data for analyses. This study proposed developing a data warehouse to analyze data in social media such as likes, comments, and sentiment. Since the sentiment is not provided directly from social media data, this study uses lexicon based classification to categorize the sentiment of users' comments. This data warehouse provides business intelligence to see the performance of the company based on their social media data. The data warehouse is built using three travel companies in Indonesia. As a result, this data warehouse provides the comparison of the performance based on the social media data.

KEYWORDS Business intelligence, lexicon based classification, sentiment analysis, social media

---

1. INTRODUCTION

The development of air transportation and airlines in Indonesia is increasing. This is marked by the growing number of airlines that have sprung up by offering both domestic and international travel routes that make the competition more competitive. With competitive competition, many airlines offer promotions that can be an attraction for consumers. This is certainly a great opportunity for business people to use information technology. The development of telecommunication and computer technology led to changes in the pattern of instant purchasing, online reservations, and the ticketing process, which in the aviation world is often called the online system or electronic ticketing (Atmadjati, 2012).

In Indonesia electronic ticketing providers are becoming more common, so competition is increasing. Because business competition requires price matching, companies must compete to attract consumers as much as possible in order to survive. Many companies use media for marketing. This includes social media, like Facebook and Twitter. With social media, customers can easily contact the company (customer service). Businesses start looking at such technologies as effective mechanisms to interact more with their customers (Ali Abdallah Alalwan, et al. 2017).

Social media has become the largest data source of public opinion (Shuyuan Deng, 2017).

Indonesia has the fourth most Facebook users in the world. Therefore, this study focuses on the relationship of social media use, namely Facebook and Twitter, to see the interaction between companies and consumers.

Data that exist in social media can help us to do the analysis to help companies get feedback from consumers. The data that can be retrieved include "like, comment, and share" information. Sentiment analysis can be used to process comments in order to get feedback on the nature of the comment, good or bad (He, Zha, & Li, 2013). Poor comments can be used as advice and input for the company in the future (Saragih & Girsang, 2017).

In this study, using existing data in social media Facebook and Twitter is expected to create business intelligence that can help analyze travel business companies in Indonesia with social media data interaction.

2. CONCEPTUAL BACKGROUND

In this chapter, we examine the concept and characteristics of business intelligence and sentiment analysis using lexicon based classification.

2.1 Business Intelligence

Business information and business analysis in the context of business processes are the key that leads to decision-making and actions that lead to improved business performance. Business intelligence can be defined as “a set of mathematical models and analytical methodologies used to exploit the data available to produce information and knowledge useful for complex decision-making processes” (Vercellis,2006, Williams, S., and Williams, N, 2006).

Advantages of business intelligence:

- Effective decisions: Business intelligence applications allow users to use more reliable information and knowledge. The result is a decision maker can make better decisions and match goals with the help of business intelligence.
- Timely decision: Dynamic, where decisions can be taken quickly. The result obtained by the organization is that the organization will have the ability to react continuously in accordance with the movements of competitors and to change when there are important new market circumstances.
- Increase Profits: Business intelligence can help business clients to evaluate customer value and desire for short-term profits and to use the knowledge used to differentiate between profitable customers and non-profitable customers.
- Reduced costs: Reducing the investment needed to use sales, business intelligence can be used to assist in evaluating the organization's costs.
- Develop Customer Relationship Management (CRM): This is essentially a business intelligence application that applies customer information collection analysis to provide responsible customer service responsibilities that have been developed.
- Reduce the Risk: Applying the business intelligence method to enter data can develop a credit risk analysis, looking at the analysis of consumer activity, producers, and reliability can provide insight into how to shorten the supply chain

2.2 Sentiment Analysis

Sentiment analysis or opinion mining is a process of understanding, extracting and processing textual data automatically to get sentiment information contained in an opinion sentence. Sentiment analysis is done to view opinions or opinion tendency of a problem or object by someone. Sentiment analysis can be distinguished based on the data source, some of the level that is often used in research sentiment analysis is sentiment analysis at document level and sentiment analysis at sentence level (Bo,P et al. 2002)

The lexicon-based approach depends on the words in the opinion (sentiment), specifically words that usually expresses a positive sentiment or negative sentiment. Words that describe the desired state (e.g. great, good) have positive polarity, whereas the words describing the unwanted state have negative polarity (e.g. bad, horrible). One common approach used in performing sentiment analysis is using a dictionary based approach. Because this research is based on Indonesia, the dictionary will use Indonesian words. Figure 1 is a positive dictionary and Figure 2 is a negative dictionary.

3. METHODOLOGY

abadi	ajek	alistis
abid	ajojing	altruistis
abrar	ajukan	amal
absolut	akademisi	amalan
acuh	akal	aman
acuhan	akas	aman
acung	akbar	amanah
adab	akhlak	amanat

Figure 1 Positive dictionary.

abaikan	alig	anti
abab	alih	antisosial
ababil	aliheo	anyang
abai	alkoholik	anyep
abal	alot	anyih
abanggribang	alpa	anyik
abn	alter	apak
abnormal	amang	apatis
abol	amarah	apes
aborsi	amatiran	apeu
abrik	amborong	arak

Figure 2 Negative dictionary.

Research conducted begins based on the interest of the writer about the data that exist on social media.

Therefore, through this research, the author wants to create a data warehouse for social media data in order to perform analyses related to social media interactions. These include an analysis of how actively the company replies or communicates with its customers on social media such as Facebook or Twitter.

3.1 Crawling Data

Data retrieval is done from selected social media platforms such as Facebook and Twitter via the social media API available on each platform.

Data retrieval is done periodically by crawlers. The data is taken every Wednesday and Saturday. This is done because the data provided by the Twitter API only retrieves data up to seven days old. For example, data retrieved on October 18, 2017 from Twitter can only go back as early as October 11, 2017. Data before that date cannot be retrieved.

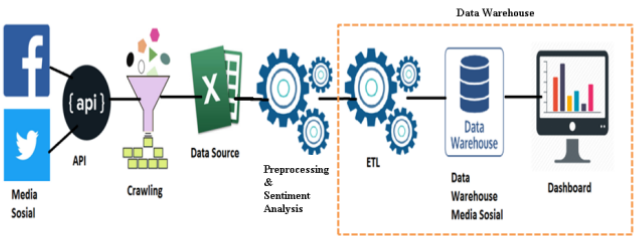


Figure 3 Methodology

From the data that was regularly taken by the crawler, was stored on in the form of excel files.

The types of data stored on each social media platform are different:

- Facebook: post, comment, reply, like
- Twitter: tweet, retweet, mention

Crawling data in this research uses Rstudio, for crawling Facebook the Library Rfacebook was used and for Twitter, TwitterR was used.

3.1.1 Crawling Facebook

In this research, will use three months of data, from September 2017 to December 2017 from three companies. The pseudocode used to get data using Rfacebook in Rstudio was:

- Load Rfacebook
- Connect to Facebook API using fbOauth
- Get Paget from Official Facebook Page using function GetPage
- Get all post in Page use GetPost
- Get Like and Comment from Post (post$Likes & post$Comments)
- Get Like and Reply from Comment using getCommentReplies
- Export to Csv format

3.1.2 Crawling Twitter

TwitterR uses the Twitter API to get the data. Because of this, there is a seven day limitation from the day we request data. The pseudocode to get the data using TwitterR in Rstudio was:

- Load TwitterR
- Connect to Twitter API using setup_twitter_oauth
- Search @from Twitter@ example from:traveloka
- Search “@” example @traveloka
- Search “to” tweet example to:traveloka
- Export to csv format

3.2 Sentiment Analysis

3.2.1 Preprocessing

Preprocessing data data comments from Facebook and Twitter social media is done by preprocessing before sentiment analysis. Figure 4 shows the preprocessing stages.

The first step is case folding. Case folding is the process of converting words into lowercase. The purpose of turning words into lowercase is to eliminate case sensitive errors. The next step is to filter the sentence. Written words are

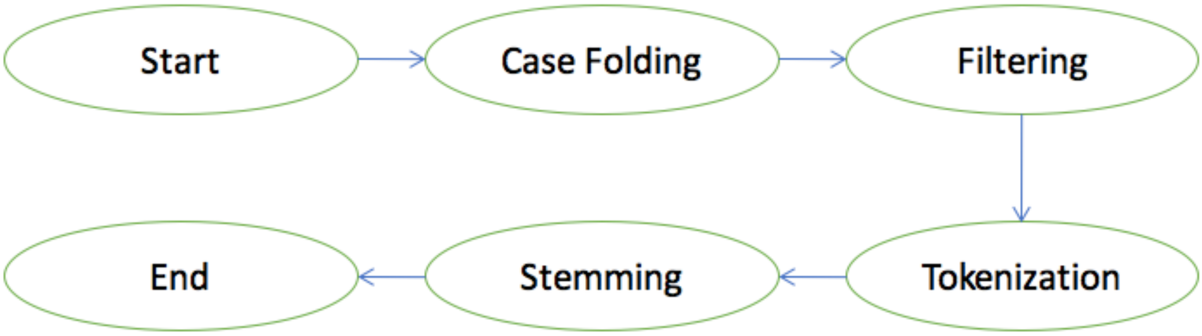


Figure 4 Preprocessing stages.

punctuation, number, and website address. The process of separating sentences into individual words is usually called tokenization. The easiest way to turn a sentence into words is to separate them with spaces. Stemming is the process of converting words into basic words.

3.2.2 Lexicon Based Algorithm

The lexicon algorithm converts data via a function that will process every sentence in the data source. Figure 5 is the pseudocode for the sentiment analysis using the lexicon based algorithm (Chopra and Bhatia, 2016).

1. Enter the text as input.

2. Divide this paragraph into tokens and store the words in an array list.
3. Select the first word from array list.

4. Fetch the words of database in second array named as database array.
5. Check whether selected paragraph word matched with each word of database array.
(i) If match found

(a) Find the sentiment of word from database whether it is positive/negative or neutral.

(b) Find the exact position of word in the paragraph.

(c) Highlight the word according to their sentiment; make it green if it is positive, red if it is negative and blue if it is neutral.
(d) Calculate the score of sentence.

(e) Store the results in database.

(ii) Else match not found

(a) Select next word from the array

(b) Go to step 5.

6. Display the result to the user.

7. Plot the graph according to the results.

Figure 5 Pseudocode for the sentiment analysis using the lexicon based algorithm.

4. RESULTS

Result from the methodology above are shown in Figure 6. There are two table facts and five dimension tables. The two fact tables are: the fact company activity and fact user activity. The five dimension tables are: dim user, dim sentiment, dim company, dim media social, and dim time.

Dashboard admin activity consists of four reports (Figure 7). The first report is the report of admin activity trends during the month, the second report provides an overview of the activities undertaken by the admin, the third report is a report of activity per day while the latter is an hourly activity. Uniquely by using the business intelligence program tableau all existing reports can affect each other, for example when we click on the first report graph on the line Traveloka and September all reports on this page will show Facebook Traveloka data in September.

Dashboard user activity consists of five reports (Figure 8). The first report is the report of user activity trends during the month, the second report is sentiment analysis report, the third report is the most active user in social media, the fourth is user activity by day and the last is an activity report by hour. With this dashboard we can analyze who is active during the month or day or time we choose in the dashboard.

On the dashboard the activity of the companies assesed can be seen. Facebook social media shows that the company Pegi Pegi is the most active compared to other companies. In September it was found that Pegi-Pegi made a ocial media strategy change, which can be seen in October with a rise of almost 368.81%. The company, Ticket, had the lowest activity. In this company there is even a decline in October and December.

On Twitter, Traveloka has the most activity compared to other companies. Traveloka has more than 1,000 activities per month. Other companies have almost 10 times less activity than Traveloka. Pegi-Pegi and Ticket had an

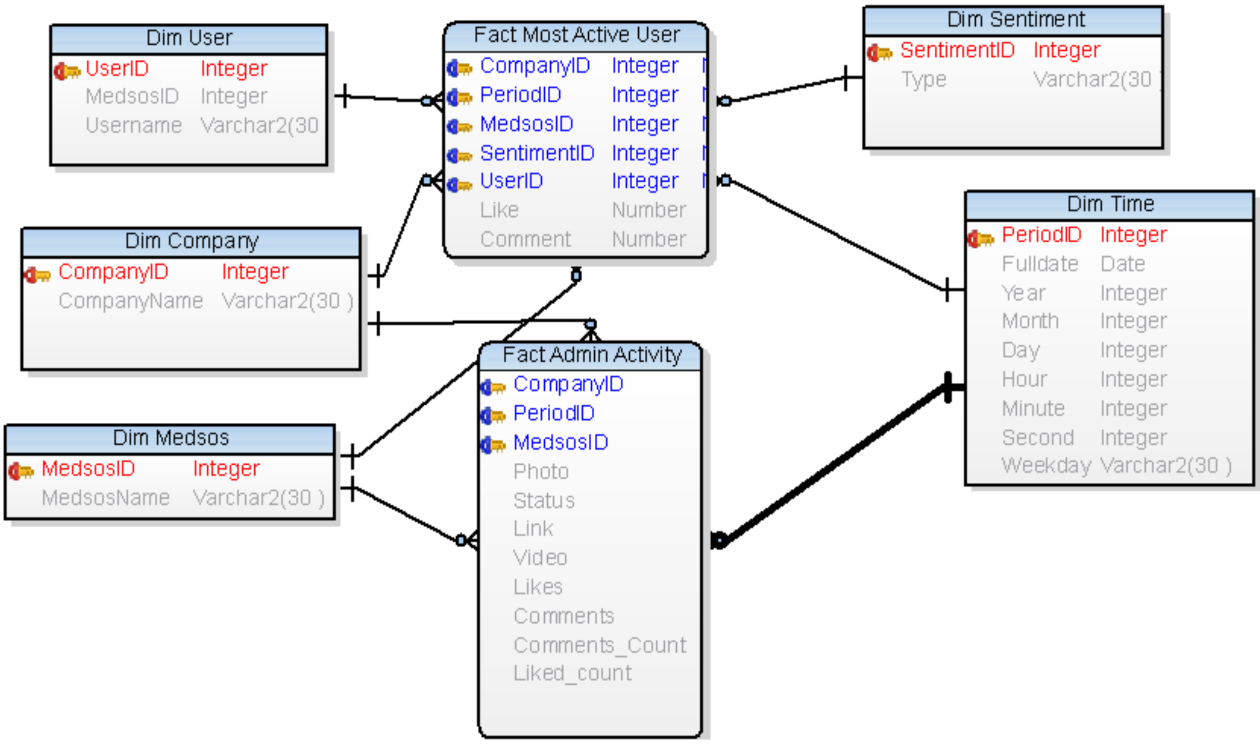


Figure 6 Star schema.

increase in November and December. In November there was a decrease in activity. Figure 9 summarizes the company activity on social media.

The most frequent Facebook activity by companies is reply to comments from customers. This was most frequently done by Traveloka, followed by Pegi-Pegi and Ticket. At Pegi-Pegi the most most common activity was liking comments from its customers. Figure 10 shows activity by hour.

The companies' Facebook and Twitter activity peaked at 16:00-16:59. Traveloka's activity peaked at 19.00 - 19.59 while Pegi-Pegi was most active at 16.00 - 16.59 and Ticket was most active at 12.00 - 12.59 (Figure 11).

Research conducted during four months of social media data collection on Facebook and Twitter, obtained 28,445 comments and

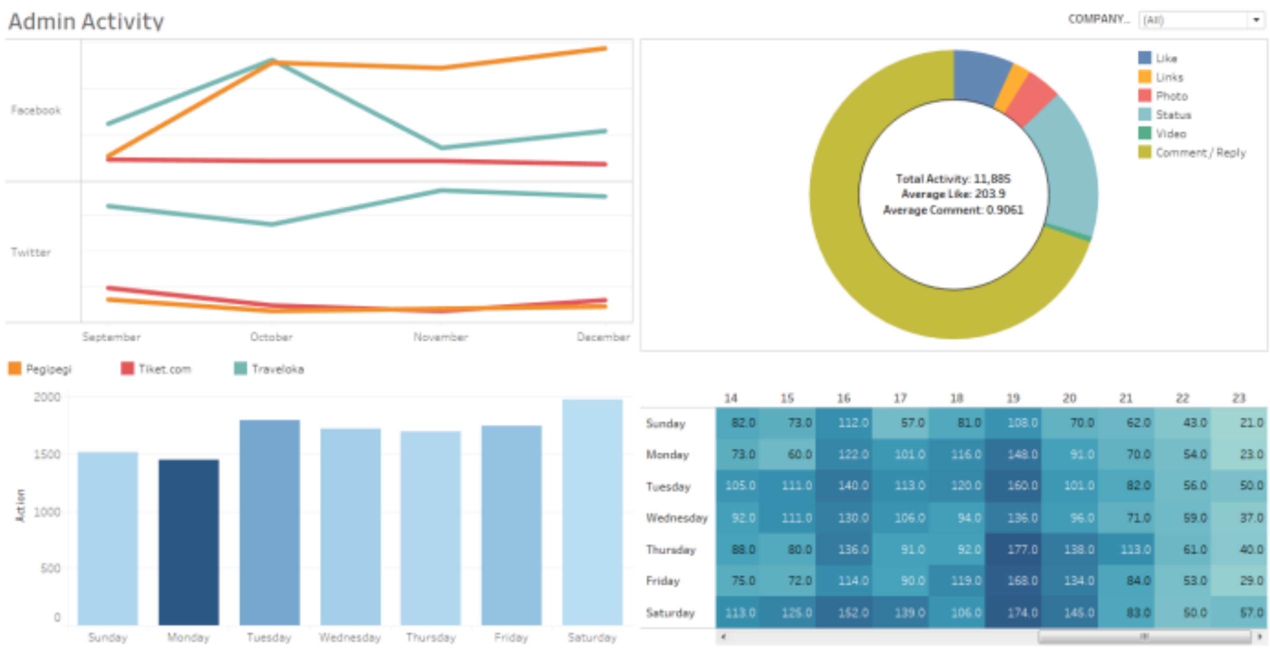


Figure 7 Dashboard company activity.

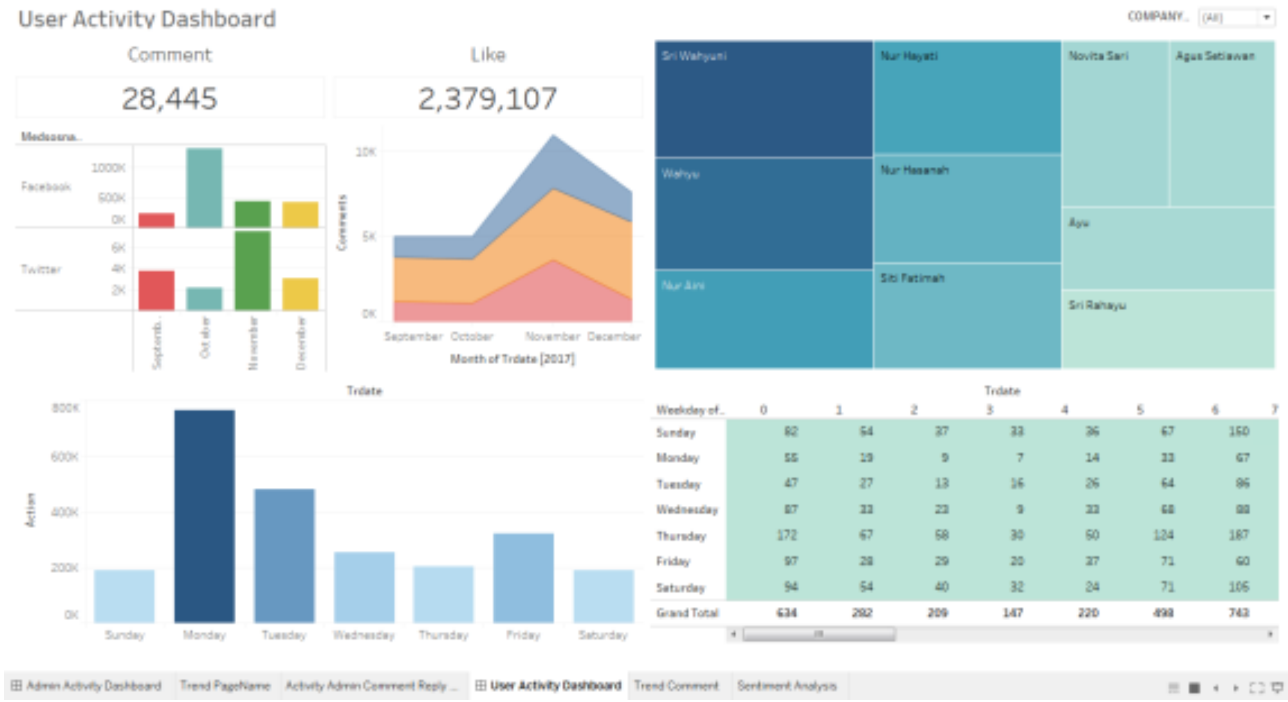


Figure 8 Dashboard user activity.

2,379,107 liked statuses by the users (Figure 12). This figure is very high, and reflects how enthusiastic the users with activities performed by the company. On social media Facebook, Traveloka has more enthusiastic users than the other two companies, this is evidenced by the existence of 1,386,318 user activity data points, of which 942,769 activities occurred in October. When viewed in more detail, Pegi-Pegi has more active users than Traveloka in the last two months (November and December).

From 28,445 comments, Traveloka has the most negative sentiment with an average of 14.26% negative, 34.51% positive sentiment and 51.23% neutral sentiment on Twitter. Tickets have the best positive analytical sentiment with a value of 44.05%, compared with negative sentiment which is only 14.10% and a neutral value of 41.85%. Figure 13 shows the results of the lexicon-based sentiment analysis.

The last four months' data got the names of users who most actively made comments or liked a status or comment. In every form of social media there were users who engaged in more than 100 activities in the last 4 months (Figure 14). On Traveloka, the top ten people engaging had an average activity of 200 interactions, while Pegi-Pegi had an average of 168 activities and Ticket has the lowest average of 84.

		Medsosname			
		Facebook		Twitter	
Pegipegi	September 2017	109		323	
	October 2017	511	368.81%	162	-49.85%
	November 2017	488	-2.28%	194	9.43%
	December 2017	573	5.50%	226	5.22%
Tiket.com	September 2017	94		485	
	October 2017	88	-6.38%	240	-50.52%
	November 2017	88	0.00%	162	-17.84%
	December 2017	74	-5.61%	314	24.68%
Traveloka	September 2017	248		1,631	
	October 2017	523	110.89%	1,374	-15.76%
	November 2017	144	-47.53%	1,851	16.07%
	December 2017	217	14.65%	1,766	-1.55%

Figure 9 Summary company activity.

5. RECOMMENDIATION

From the dashboard analysis various recommendations for companies studied were obtained.

5.1 Traveloka

On Facebook social media needs to be improved again because from November there was a significant decline (23%) compared to the previous month. At 19.00 - 19.59 the activities of the Traveloka are recommended to have more human resources in order to help solve customer problems.

	Facebook					
COMPANYN..	Like	Links	Photo	Status	Video	Comment / Reply
Pegipegi	726.0	150.0	153.0	2.0	7.0	643.0
Tiket.com	2.0	93.0	185.0	1.0	61.0	2.0
Traveloka	77.0	1.0	131.0	1.0	7.0	915.0

Figure 10 Detail company activity.

	Facebook			Twitter		
	Pegipegi	Tiket.co..	Travelo..	Pegipegi	Tiket.co..	Travelo..
0	20.0	1.0	22.0	14.0	6.0	139.0
1	7.0		9.0	12.0	3.0	52.0
2	5.0		6.0	1.0	3.0	61.0
3	2.0		6.0	3.0	4.0	29.0
4	6.0		6.0	1.0	1.0	37.0
5	23.0		14.0	1.0	7.0	96.0
6	75.0		20.0	5.0	16.0	153.0
7	63.0		33.0	18.0	9.0	186.0
8	73.0	24.0	46.0	28.0	138.0	315.0
9	73.0	1.0	56.0	41.0	62.0	380.0
10	172.0	3.0	67.0	54.0	81.0	361.0
11	141.0	6.0	47.0	140.0	42.0	347.0
12	126.0	100.0	49.0	60.0	102.0	357.0
13	105.0	5.0	42.0	41.0	24.0	366.0
14	90.0	7.0	38.0	53.0	113.0	327.0
15	123.0	2.0	61.0	38.0	44.0	364.0
16	155.0	92.0	38.0	153.0	110.0	358.0
17	118.0	15.0	37.0	35.0	72.0	420.0
18	58.0	5.0	100.0	24.0	129.0	412.0
19	99.0	5.0	171.0	136.0	56.0	604.0
20	41.0	73.0	102.0	27.0	101.0	431.0
21	47.0	3.0	68.0	13.0	35.0	399.0
22	29.0	1.0	55.0	4.0	25.0	262.0
23	30.0	1.0	39.0	3.0	18.0	166.0

Figure 11 Detail company activity in hour.

5.2 Ticket

On Facebook, social media needs to be improved. In September there were 94 activities, but this declined considerably to 74 activities in December. On Twitter, engagement should be improved again as compared to Traveloka, as the activity of Ticket is lagging behind. For Twitter we suggest human resources should be available in the early hours, as in December at 00.00 - 07.00 there are only seven activities, compared with

Medsosna..	COMPANYN..	September	October	November	December	Grand Total
Facebook	Pegipegi	475	358,240	285,131	348,364	992,210
	Tiket.com	4,940	3,597	1,707	2,362	12,606
	Traveloka	233,422	942,769	144,366	65,761	1,386,318
Twitter	Pegipegi	787	126	168	130	1,211
	Tiket.com	832	410	438	740	2,420
	Traveloka	2,155	1,604	6,894	2,134	12,787

Figure 12 Summary user activity.

user activity on Ticket's Twitter feed of as much as 85 activities.

5.3 Pegi-Pegi

For Twitter, we suggest increased human resources in early hours. In December at 00.00 - 07.00 there were 55 activities only compared with user activity on Twitter Pegi - Pegi as many as 244 activities.

6. CONCLUSION

Based on the results of the research, there are several conclusions. By using business intelligence conducted in this research, Traveloka has the most interaction in social media, as compared with Pegi-Pegi and Ticket.com.

This research provides some suggestions for the development of business intelligence for social media interaction. The classification accuracy can be further improved by using algorithms and machine learning such as naive baise classification and in the future data could also be analyzed to include emoticons for more complete information from Facebook.

			Month of Trdate									
			September 2017		October 2017		November 2017		December 2017		Grand Total	
COMPANYN..	Medsosna..	Description	Comme..	% of Total C..	Comme..	% of Total C..	Comme..	% of Total C..	Comme..	% of Total C..	Comme..	% of Total C..
Pegipegi	Facebook	Negative			57	8.32%	141	9.26%	402	10.58%	600	9.97%
		Neutral	5	45.45%	477	69.64%	1,045	68.66%	2,513	66.17%	4,040	67.15%
		Positive	6	54.55%	151	22.04%	336	22.08%	883	23.25%	1,376	22.87%
	Twitter	Negative	88	11.18%	10	7.94%	35	20.83%	20	15.38%	153	12.63%
		Neutral	344	43.71%	52	41.27%	65	38.69%	48	36.92%	509	42.03%
		Positive	355	45.11%	64	50.79%	68	40.48%	62	47.69%	549	45.33%
Tiket.com	Facebook	Negative	10	7.04%	2	5.71%	3	9.68%	1	1.04%	16	5.26%
		Neutral	87	61.27%	28	80.00%	21	67.74%	74	77.08%	210	69.08%
		Positive	45	31.69%	5	14.29%	7	22.58%	21	21.88%	78	25.66%
	Twitter	Negative	116	13.94%	78	19.02%	75	17.12%	99	13.38%	368	15.21%
		Neutral	333	40.02%	121	29.51%	173	39.50%	303	40.95%	930	38.43%
		Positive	383	46.03%	211	51.46%	190	43.38%	338	45.68%	1,122	46.36%
Traveloka	Facebook	Negative	59	5.64%	168	7.95%	246	13.18%	75	11.01%	548	9.60%
		Neutral	679	64.91%	1,202	56.86%	800	42.87%	411	60.35%	3,092	54.18%
		Positive	308	29.45%	744	35.19%	820	43.94%	195	28.63%	2,067	36.22%
	Twitter	Negative	344	15.96%	342	21.32%	1,053	15.27%	350	16.40%	2,089	16.34%
		Neutral	871	40.42%	584	36.41%	4,064	58.95%	864	40.49%	6,383	49.92%
		Positive	940	43.62%	678	42.27%	1,777	25.78%	920	43.11%	4,315	33.75%

Figure 13 Sentiment analysis.

Username	Traveloka	Username	Tiket.com	Username	Pegipegi
Wahyu	273.0	Neo Radhita	171.0	Maresa Sumardi	236.00
Sri Wahyuni	250.0	TraxFMJKT	127.0	Nur Hasanah	202.00
Siti Fatimah	231.0	Cassiopiea Abdullah	108.0	Arjuna Ireng	163.00
Ayu	206.0	Maman Hasan	98.0	Nur Halimah	161.00
Wulan Dari	200.0	Benny Kautsar	78.0	Rian	158.00
Nur Hayati	180.0	Samuel Dm	78.0	Tanty Phoa	155.00
Novita Sari	178.0	Donny Achmad Fauzi	53.0	Agus Setiawan	154.00
Irfan	177.0	Nur Aini	44.0	Sri Rahayu	154.00
Putra	170.0	Reza Hafidz Hafidz	43.0	Nur Hayati	152.00
Rizal	167.0	Linda	39.0	John Hendry Andri	149.00
Average	203.2	Average	83.9	Average	168.40

Figure 14 Most active users.

7. REFERENCES

Adriani, M., Asian, J., Nazief, B., Tahaghoghi, S. M., & Williams, H. E. (2007). Stemming Indonesian: A confix-stripping approach. *Journal ACM Transactions on Asian Language Information Processing (TALIP).*

Alalwan, A. A., Rana, N. P., Dwivedi, Y. K., & Algharabatc, R. (2017). Social media in marketing: A review and analysis of the existing literature. *Telematics and Informatics*, 1177-1190.

Atmadjati, A. (2012). *Era Maskapai Saat Ini.* Yogyakarta: Leutika Prio.

Barlow, J., & Maul, D. (2000). *Emotional Value: Creating Strong Bonds with Your Customers.* San Francisco: Berrett-Koehler Pub-lishers, Inc.

Bo, P., Lee, L., & Vaithyanathan, S. (2002). Thumbs up? Sentiment Classification Using Machine Learning Techniques. *EMNLP*.

Budiwati, S. D., & Setiawan, N. N. (2018). Experiment on building Sundanese lexical database based on WordNet. *Journal of Physics: Conference Series*.

Chopra, F. K., & Bhatia, R. (2016). Sentiment Analyzing by Dictionary based Approach. *International Journal of Computer Applications*, 32-34.

Deng, S., Sinha, A. P., & Zhao, H. (2017). Adapting sentiment lexicons to domain-specific social media texts. *Decision Support Systems*, 65-76.

Girsang, A. S., & Prakoso, C. W. (2017). Data Warehouse Development for Customer WIFI Access Service at a Telecommunication Company. *International Journal on Communications Antenna and Propagation.*

He, W., Zha, S., & Li, L. (2013). Social media competitive analysis and text mining: A Case study in the pizza Industry. *Internasional Journal of Information Management*, 462-472.

Moro, S., Rita, P., & Vala, B. (2016). Predicting social media performance metrics and evaluation of the impact on brand building: A data mining approach. *Journal of Business Research*, 3341-3351.

Ray, P., & Chakrabarti, A. (2017). Twitter sentiment analysis for product review using lexicon method. *International Conference on Data Management, Analytics and Innovation (ICDMAI)*, 211-216.

Saragih, M. H., & Girsang, A. S. (2017). Sentiment analysis of customer engagement on social media in transport online. *Sustainable Information Engineering and Technology (SIET)*, 24-29.

Vercellis, C. (2009). *Business Intelligence: Data Mining and Optimization for Decision Making*. Politecnico di Milano: Wiley.

Williams, S., & Williams, N. (2006). *The Profit Impact of Business Intelligence.* San Francisco: Morgan Kaufmann.