A New Era For Sustainability Assessment Of Firms? Achievements And Challenges In Incorporating Big Data And Artificial Intelligence Into Existing Frameworks

ARON BELINKY
Escola de Administração de Empresas de São Paulo (EAESP/FGV)
MARIO H. N. SARAIVA
Columbia Center for Sustainable Investment (CCSI/UoC)
ALEXANDRE MIYAKE DA SILVA
Escola de Administração de Empresas de São Paulo (EAESP/FGV)

Resumo

Assessing the performance, risks and opportunities of companies from a sustainability perspective is a matter of growing interest to managers, investors and other stakeholders. However, it is also a challenging task, especially due to the combination between the complexity of the sustainable development agenda and the volume and diversity of information needed. Usual methods rely mainly on companies' disclosures and environmental, social and governance (ESG) frameworks, but two recent facts are expected to re-shape the field: the United Nations' Agenda 2030 and its Sustainable Development Goals (SDG), and the use of Big Data and Artificial Intelligence (BDAI) in the field of corporate sustainability, especially applying Natural Language Processing (NLP) to the extensively available texts containing sustainability relevant information produced by and about firms. Building mainly on a series of roundtables and interviews with experts, this article presents novel insights into the current state of BDAI applied to sustainability assessment of firms. Key conclusions include the observation that BDAI can contribute to largely improve Corporate Sustainability Assessment (CSA) tools, but also that BDAI by itself is not a standalone solution to the pressing challenges of Environmental Social and Governance (ESG) oriented investing and management. It was perceived that BDAI is likely to increase the strengths of current ESG frameworks, but also their weaknesses, as it will not resolve the underlying challenge of the lack of standardization regarding corporate ESG performance assessment. Furthermore, the current frameworks on corporate sustainability – represented by the ESG perspective embedded in common CSA tools and methods – are related to (but not the same as) the paradigm represented by Agenda 2030. Therefore, there is a need for harmonization, between ESG assessment tools and the SDGs.

Palavras chave: Sustainability, Big Data, Artificial Intelligence, ESG assessment, Sustainable Development Goals - SDG.

A New Era For Sustainability Assessment Of Firms? Achievements And Challenges In Incorporating Big Data And Artificial Intelligence Into Existing Frameworks

INTRODUCTION

This article discusses the possible role of Big Data Analytics and Artificial Intelligence (BDAI) in guiding private investment on companies, based on the assessment of their actual and/or expected contribution to sustainable development. There is a growing expectation that private investment should drive businesses towards sustainability, by prioritizing companies that contribute to a sustainable society. Current Corporate Sustainability Assessment (CSA) tools, such as traditional environmental, social and governance (ESG) ratings and analysis of sustainability reports, struggle to overcome the challenges of limited ESG data on companies and self-reporting bias. Additionally, there are no broadly agreed standards to assess the contribution of businesses towards sustainable development.

In 2015 the launch of the United Nations (UN) Sustainable Development Goals (SDGs), as part of Agenda 2030, articulated and updated the global sustainable development agenda, setting interdependent goals and targets that encompass and go beyond existing ESG perspectives, emphasizing integration and interdependence among the various issues, as well as the scale, urgency, and inclusiveness of the changes needed (UN, 2015). Although Agenda 2030 and its SDGs were created primarily under the perspective of national governments, they also provide a powerful framework for the private sector and bring into question some aspects of existing ESG concepts, frameworks and tools.

In such context, this article presents a discussion on the ways of measuring firm's contribution (either positive or negative) on ESG factors, using BDAI to improve the assessment capacity on sustainability performance.

In addition to desk research, we have gathered several experts from different countries and backgrounds in two different roundtable discussions – in New York City and in São Paulo – to discuss those topics. Out of the experts from 16 countries, we had participants from management and sustainability academia, data science, business, non-governmental organizations, standards setters and government. Research was made in February-September 2018, and this paper presents some of the main insights concerning the actual and potential use of BDAI for CSA.

At both roundtables, experts agreed that BDAI can contribute to largely improve CSA tools, but also that BDAI by itself is not a standalone solution to the pressing challenges of ESG investing and management. It was perceived that BDAI is likely to increase the strengths of current ESG frameworks, but also their weaknesses, as it will not resolve the underlying challenge of the lack of standardization regarding corporate ESG performance assessment. Furthermore, the current frameworks on corporate sustainability – represented by the ESG perspective embedded in common CSA tools and methods – are related to (but not the same as) the paradigm represented by Agenda 2030 and its SDGs. Therefore, there is a need for harmonization, between ESG assessment tools and the SDGs. An important issue, then, lies in the incongruences between ESG assessment tools, which are likely to escalate as they try to incorporate the SDG framework and increase the use of BDAI.

Another important conclusion is that the creation of new CSA tools specifically from the SDG perspective should be avoided, in order not to increase the multiplicity of instruments and the reporting burden on companies. It is also important to emphasize that the need for improvement in existing ESG assessment tools is often highlighted by stakeholders,

including investors, who report that existing methodologies are not sufficient to their decision-making needs (BlackRock, 2016; Amel-Zadeh, 2017).

It is important to underline that the focus of the debates was on extracting information from reports and public information produced by and about companies in the format of natural language, such as reports, news and other common publications. The possibility of exploring other sources commonly used in BDAI applications - such as the data automatically produced by transactions and sensors - for CSA was mentioned as a field of huge potential, but out of the scope of this research.

This paper is structured as follows: first, the context for the three main topics of the paper - Disclosure for Sustainable Investment, Corporate Sustainability Assessment (CSA) Frameworks and Big Data and Artificial Intelligence for CSA- is set, then the research methodology is further explained, followed by our findings, conclusions and recommendations to future research.

Disclosure for Sustainable Investment

It is becoming clear that companies able to create long-term shared value for investors, stakeholders and society while contributing to sustainable development and respecting planetary boundaries will have better chances to succeed in the emerging business environment (Porter & Kramer, 2006, 2011; Belinky, 2017; Vendramini & Belinky, 2017). Identifying such companies is becoming a priority, bringing together the issues of corporate transparency, reporting and assessment, and the fields of Corporate Social Responsibility (CSR), Corporate Sustainability and Sustainable Development, among others.

A growing number of investors have understood the broad impact of their investments and are now requiring not only financial returns but also that a company's business practices are mitigating negative ESG impacts and risks while maximizing societal benefits (BlackRock, 2016; US|SIF, 2016). In the last years, increasing demands for ESG data by a range of different stakeholders – such as investors, managers, governments, consumers and civil society organizations – have encouraged research companies on financial data, as well as specialized ESG information suppliers (ESG-IS), to develop new business streams.

Stakeholder theory is commonly used to explain the increasing global trend in CSR and, consequently, sustainability reporting (Nielsen & Thomsen, 2007; Russo & Perrini, 2010; Liang & Renneboog, 2017). In a way, as society provides corporations legitimacy and a license to operate, companies are expected to conduct their business in a socially responsible manner. Then, the demand for sustainability reporting can be seen as an expression of the general belief that companies must conduct themselves responsibly. Transparency through reporting becomes a strategic issue to companies, necessary to preserve business legitimacy. This growing trend towards ESG disclosure and investing creates a demand for ESG data on companies. Table 1 displays some examples of ESG data providers.

[Table 1]

In response to the demand for better disclosure from companies on sustainability practices, the Global Reporting Initiative (GRI) developed comprehensive guidelines to help businesses better disclose information that is relevant to stakeholders (Brown, Jong, & Lessidrenska, 2009; Lozano & Huisingh, 2011). In response to the lack of official or self-regulatory standards and guidelines, GRI has been a key player in the field of voluntary sustainability reporting. GRI guidelines are followed by thousands of companies worldwide, being the most widely used guidelines for sustainability reporting (Bradford, Earp, Showalter, & Williams, 2017). However, a trade-off of GRI's broad applicability and flexibility is the

difficulty to establish systematic comparability schemes, that could support automatic analysis and assurance or auditing procedures.

In 2011, the Sustainability Accounting Standards Board (SASB) focused their efforts on understanding what is material to investors and developed an industry-specific materiality map. According to SASB (2018), companies do not need to report on every ESG issue but it is their fiduciary duty to report "any known trends or uncertainties that have had or that the registrant reasonably expects will have a material favorable or unfavorable impact on net sales or revenues or income from continuing operations." According to some Brazilian companies interviewed (Saraiva, 2017), SASB has been recognized as an important milestone in sustainability accounting but it is criticized for being US-centric in their approach.

There is also a move towards Integrated Reporting - a reporting framework aiming to provide objective and understandable information on companies regarding value generation to different stakeholders. The initiative developed by the International Integrated Reporting Council (IIRC) is a first attempt to integrate mainstream business methods such as financial reports with non-financial ones, such as sustainability reports (IIRC, 2013). It has developed conceptual and practical guidance on how to consider other capitals related to a firm's health and performance (beyond economical and financial). Integrated Reporting is committed to build a system that can address and clearly report the actual factors that support a company's long-term success.

As it has been widely discussed, the field of sustainability reporting lacks broad and internationally adopted standards such as the Generally Accepted Accounting Principles (GAAP) for financial accounting. In many countries, GAAP is mandatory while sustainability reporting remains voluntary (Tschopp & Huefner, 2015).

Although the idea of making sustainability reporting mandatory or regulated have been mentioned and even implemented in different contexts, there is broad consensus that such measures could have negative effects, mainly because of the diversity and complexity of the topics involved - they are likely to be either too complex or unduly narrow, thus creating confusion or "lowering the bar" on such a crucial topic. (ISO, 2010; 2017)

The context described above, together with a plethora of stakeholders requesting information, creates a difficult environment for sustainability teams, which are usually small, (Searcy & Buslovich, 2014) to produce a comprehensive yet short sustainability report (another conundrum of reporting). This also poses potential implications on the comparability of reports (Morsing, Schultz, & Nielsen 2008; Dingwerth & Eichinger 2010; Bradford, Showalter, & Williams, 2017), impairing the analysis of stakeholders, especially investors, not to mention companies' own management. Arvidsson (2014) suggests that the problem lies on the fact that financial analysts at international investment banks have "severe problems with understanding CSR information per se and how to use it in a valuation context".

Based on our research, it seems that the actual usability of sustainability reports is still limited in great measure due to the lack of (i) systematic and objectively comparable indicators and, in some cases, (ii) credibility. The first aspect is related to the standardization issue, which does not seem to be near to a harmonization. Regarding credibility, audit processes are cited as a possible solution that could "reduce some of the skepticism and also improve the quality and comparability of CSR information" (Arvidsson, 2014). Although there are already many sustainability reports that are subject to some level of assurance, the problem lies again on the fact that, for the very nature of the sustainability field, it is not likely that rigid mandatory rules, dependent on standardization and enforcement, would work well.

As discussed further in this article, we believe that the SDGs and BDAI are potential drivers to increase the quality and usability of CSA.

Agenda 2030 and its Sustainable Development Goals

On September 25, 2015, the United Nations General Assembly adopted the Agenda 2030 for Sustainable Development, which stated that "this Agenda is a plan of action for people, planet and prosperity. It also seeks to strengthen universal peace in larger freedom" (UN, 2015). In a way, Agenda 2030 organizes the sustainable development agenda globally, with broad support from governments, companies and society in general, as represented at the UN.

Based on historical adjournment and integration of the international commitments for sustainable development represented by Agenda 2030, the SDGs bring a common language between governments, businesses, and citizens to discuss sustainability at the global level. By facilitating communication between different actors, the SDGs present an opportunity to advance the harmonization of CSA tools and contribute towards the discussion on an internationally accepted sustainability reporting and assessment referential that would facilitate sustainable investments.

In particular, the SDGs put emphasis on integration and interdependence among the various goals, as well as on the scale and urgency of the changes needed. These issues are often overlooked by existing ESG frameworks. As in a natural ecosystem, integration and interdependence are important characteristics of a sustainable society. For example, the water from streams used to irrigate crops is also an important natural source for the local fauna and it could have generational cultural value for local communities. From this perspective, it does not make sense to increase food production to combat hunger while depleting local streams.

Furthermore, the SDGs are an easy-to-understand framework that can be used in moving investors from traditional to sustainable investing. It is possible that some investors might have avoided sustainable investing given the lack of clarity on the importance of their private investments towards the pressing global issues we face today.

Identifying global issues is not a challenge today as it used to be, but assessing which companies are contributing to the SDGs and which are not remains a challenge. Sustainability is a complex topic that overlaps many disciplines - natural, social, and hard sciences, and companies differ in several ways (industry, sector, size, to name a few). As mentioned, the SASB materiality map has been instrumental in identifying what topics are material in each sector. Recently, SASB has created a link between their framework and the SDGs. According to Eccles & Consolandi (2018), "a relationship between ESG outcomes and SDG impacts exists via the concept of materiality". The concept of materiality on ESG factors partially overlaps with the SDGs. But the SDGs have a much broader focus than the financial impacts of ESG factors. Furthermore, by focusing mostly on what is material to companies, they risk to underplay the interconnectedness of SDGs.

Reporting is often "used by companies in a strategic and instrumental manner" (Nielsen, & Thomsen, 2007), tending to focus on specific ESG topics - and now, on SDGs as well - that are more relevant to their priority-stakeholders in order to achieve specific interests. In doing a disentangled approach to assessing the ESG impacts, companies continue to focus on issues related to their strategic corporate goals and diverge from the interdependence of sustainability issues posed by the Agenda 2030 (Irvine, 2016).

It is also important to notice that SDG 12 drew heightened attention to corporate reporting as an important indicator of Responsible Production and Consumption. Indicator 12.6 states that one of its targets is to "encourage companies, especially large and transnational ones, to adopt sustainable practices and to integrate sustainability information into their reporting cycle" (UN, 2015).

Private companies, and especially big multinationals, have the potential to advance this Agenda, taking into account their resources and power of influence. However, there is little data that can be used to measure the contribution of companies in fulfilling Agenda

2030, and the absence of a robust framework for the link between ESG and SDG performance makes such assessment even more difficult. The fact that the SDGs were conceived as a global-level and nation-level framework to guide countries policies towards sustainability poses challenges for their adoption by the private sector. Even though there is already some effort in such direction, e.g. the *SDG Compass - The guide for business action on the SDGs* (GRI, UNGP & WBCSD, 2015), it remains unclear if mapping the existing ESG frameworks to the SDGs is enough to assess firm's contribution to such goals.

Big Data and Artificial Intelligence for Corporate Sustainability Assessment (CSA)

The advent of the internet and the evolution of Information Technologies (IT) in the last decades created an exponential increase in the production and circulation of natural language (NL) data - text, image, audio and video documents. In parallel, other two kinds of data sources are also increasingly available: those generated by electronic transactions - such as online payments, telephone calls and web interactions - and sensors - devices that automatically register and transmit data from and about almost any kind of thing or phenomenon. The enormous amount of data continuously created and circulated on the web has some common characteristics, usually referred to as the "Vs" (volume, variety, velocity and many others, depending on the author) that, together with the tools to deal with them, turns a lot of data into the field of *big data analytics*, or just *Big Data* (Borne, 2014; Francisco, 2015).

To make sense out this ocean of records, turning data into useful information, a new field of knowledge is evolving, combining powerful processing power and analytical techniques, such as machine learning and Artificial Intelligence (AI). Merriam-Webster dictionary provides a simple definition of AI, as 1) a branch of computer science dealing with the simulation of intelligent behavior in computers, and 2) the capability of a machine to imitate intelligent human behavior. The term is frequently associated with teaching computers speech recognition, visual perception, classification problems, among others. In simpler terms, computers are trained to replicate tasks performed by humans such as the ability to classify news articles into themes based on each article's content.

Exploring these phenomena and their many implications, challenges, possibilities, risks and opportunities is out of the scope of this article. Therefore, under our research findings section, we will highlight only the aspects of particular relevance to our research, from the combined perspective of the application of BDAI in the field of CSA. In short, we are interested in the organizations' novel ability to process - within reasonable limits of time and costs - such large volumes of language and numerical information. We foresee that this ability have the potential to radically improve both quantity and quality of sustainability information on companies, as well as to assist investors (and other stakeholders) in maximizing their returns (or any other expected outcomes) given their financial and non-financial preferences. Our goal in this article is not to exhaust the topic, but rather identify and report what we have found as the most present and promising areas of application of BDAI to the transparency and assessment of corporate sustainability performance.

Methodology

In view of such context, this article presents the key insights from our work to understand the state of the art about assessing firm's performance on sustainability using BDAI, as well as the possible integration of Agenda 2030 and its SDG framework into existing methods and tools. This article presents a synthesis of our main conclusions, focusing on the question "How can BDAI contribute to drive private investment on companies, based on assessment

and transparency regarding their actual and/or expected contribution to sustainable development?"

The work plan of the research - performed between March and September 2018 - included desktop research, two roundtable discussions, and one open session where, overall, experts from 36 organizations, representing 16 countries, participated.

The main objectives of the first roundtable were to (i) discuss the current context on ESG and sustainability reporting, (ii) its relevance to the various stakeholders - especially investors - and (iii) the specificities of SDG evaluation and reporting. As for the second roundtable, the focus was to (i) advance the debates about how information related to corporate sustainability - both on ESG themes and on Agenda 2030/SDGs - is currently available and processed, and (ii) evaluate the opportunities and challenges brought about by the new BDAI technologies.

RESEARCH FINDINGS

Congruence of current ESG performance measurements

ESG information suppliers (ESG-IS) have a key role in assisting investors identifying the level of a company's commitment to sustainability. This is not a simple task since sustainability is hard to measure (another reason why it has been hard to come up with standards). It is clear that there is a great diversity of methodologies on the way to measure ESG performance (Antolín-López, Delgado-Ceballos, & Montiel, 2016). In the process of trying to measure performance, each ESG-IS has come up with their own methodology. Then, in essence, all of them are trying to report on how companies are doing in terms of sustainability but their particular way of measuring such performance leads to different results.

For example, considering the sustainability scores for the top 73 large cap equities on the Alaska Retirement Management Board, Saraiva et al. (2017) found that at least 24.66% of companies had outstanding different sustainability scores. MSCI and Thomson Reuters Asset 4 scores were used in the exercise. Table 2 shows the companies with the greatest score difference between these two rating agencies.

[Table 2]

Table 2 compares two out of the many available ESG ratings. Investigating these score differences is difficult and sometimes impossible, since not every ESG-IS made available their methodology and/or original data. Several reasons could be the origin of the differences, including factors such as the different weights used for each indicator, the difference between indicators, or the time the information was captured and analyzed.

Another potential explanation is that although these two rating agencies work mainly with publicly available information on companies, it is a common practice for them to reach out to companies for further clarification, especially when a particular company is associated with controversial behavior. As already mentioned, companies are reportedly overwhelmed by many requests for information from multiple stakeholders and it is difficult for them to respond to all request.

Additionally, it is possible that some companies are aware of which rating methodology portraits them better than others. Then, given the need to prioritize disclosure of information, companies have an incentive to report more to some ratings than to others, therefore, creating information asymmetry between ESG-IS, which may lead to different ratings of the same company.

It is important to understand that these scores are subject to a strong bias if they are heavily weighted on company reports and policies rather than quantifiable and/or objective performance measures. Some companies are very good at concealing their negative impacts and/or overestimating the positive ones through advanced marketing schemes, the practice also known as greenwashing. ESG-IS consulted in our research typically try to reduce this bias by monitoring what is being said in the media about companies. If any company is caught in controversial news, ESG-IS will typically reach out for additional information.

However, although there is ESG information available in companies' reports and in the media and other publications, it is usually unstructured, in the format of texts, that many times include also non-textual information, such as images and diagrams. Of course, as previously mentioned, the quality of reports is an issue as well and we do not expect to find every ESG data from reports themselves. Information can be complemented, for instance, with information from local media that, for its bigger focus and granularity, is likely to have additional ESG information on companies. Even social networks could potentially feed the ocean of data, that also may include video and audio formats. BDAI can harness these untapped sources of data and increase the quantity and, with proper analysis, the reliability of ESG information available on firms.

Improving ESG assessment through BDAI

In the context of sustainable investments and the evolution of the ESG movement, Big Data tools and techniques significantly expanded the scope of data on companies' ESG performance, while Artificial Intelligence techniques are used to extract meaning out of these large volumes of information. For example, a simple Google search on a company's name or brand can return literally millions of results. Within reasonable costs and in a short time, with web scraping algorithms one can "download" all these results and use other algorithms to semi-structure ("clean") the data and then extract meaningful ESG information from it. Stakeholders are no longer dependent on company disclosure only, since these new technologies allow an increase on the data volumes on each company, as well as an initial quality check of the information, by triangulating information from different sources and comparing with the one disclosed.

For example, often local level newspapers and digital media have relevant ESG data. Because locals are the first to be affected by the negative impacts, they are frequently the first to spot a problem. There are several cases describing situations where disasters or chronical situations of contamination or human rights abuse that, despite being repeatedly reported in local newspapers, digital media or social networks, never reached mainstream media, or only did so after a long time. Different from human analysts, BDAI tools can search for information on several thousand sources in the exact similar manner, thus unveiling facts that otherwise won't be noticed. However, we do recognize that algorithms are not free from bias. Bias is passed on to algorithm because it is built by humans that have their own bias of the world. Nevertheless, since the tool is used uniformly to measure ESG performance, there would be less bias in comparison to different analysts analyzing and rating different companies.

For years, stakeholders that relied mainly on company disclosure through annual reports would only have a chance to react to company controversies months after the issue had happened. In accordance with the findings from literature on sustainability reporting (Searcy & Buslovich, 2014), Brazilian multinational companies reported that on average it takes six months to compile their sustainability report (Saraiva, 2017). BDAI can help sustainability teams to quickly sort through internal information and identify the most material issues to report, allowing for timely updates - quarterly, monthly, or even daily. This

close and timely monitoring of companies allows for the early identification of trends and potential crisis. For example, in 2017, while Wall Street had high earnings expectation for Facebook, TruValue Labs' Pulse Score captured a different narrative with many red flags on Facebook. By following daily updates on the company through news sources from different countries and using the SASB materiality map to look for relevant information, TruValue captured financially material ESG signals that most analysts did not.

BDAI for CSA Practical Application Challenges

The use of BDAI is perceived as a possibility to overcome some of the challenges associated with the quality and availability of sustainability information. However, it is important to remind that there are still some practical obstacles to overcome.

Many reports use visual resources such as images to disclose information rather than text or tables. Whereas humans can easily extract information from visual images, it remains a complex obstacle for machines. Image recognition algorithms are not sufficient yet to interpret infographics, such as a color-coded image describing which company goals have been achieved. There is a lack of reporting guidelines to indicate the best structure for machine-readable reports.

Natural Language Processing (NLP) is the technology to automatically extract meaning from language (texts, audios, and videos) and it relies on algorithms that have been trained to process and understand natural language-based data.

Notice that the examples described on the application of BDAI in ESG investing section were all in English. Language is another key issue in the application of BDAI to overcome ESG data gaps in non-English speaking emerging markets. The 35 emerging markets considered in UNCTAD's definition of EM, had at least 14 non-English native languages.. Only 14% of these markets have English as one of their official languages.

In specific, the building blocks of NLP are, among other things, algorithms that are used to "teach" the AI algorithms system what is relevant and what is not. These algorithms are well developed in English and we have many alternatives to perform basic NLP tasks such as parts of speech tagging, parsing in general, and textual classifications. These have been developed through years of cross-disciplinary research in linguistics and computer science.

Despite the numerous papers on text processing - sentiment analysis (Jelveh, Kogut, & Naidu, 2014), text summarization, text categorization (Jelodar, Wang, Yuan, & Feng, 2017) and entity extraction, among others - the majority of works are on the English language. The three papers (Mahadevan, Ng, & World Scientific (Firm), 2009; Liew, Adhitya, & Srinivasan, 2014; Székely & vom Brocke, 2017) on text-mining of CSR reports examined in this research excluded non-English reports. In regard to Portuguese, there are few papers exploring the application of text mining tools on Brazilian Portuguese and European Portuguese (Castro, Souza, Vitório, Santos, & Oliveira, 2017; Filho & Paulo, 2017) in comparison to text mining with English sources. This limitation on text mining tools that operate in Portuguese was further confirmed during the second roundtable, where it became clear that most algorithms were developed to operate only in English.

Nevertheless, most sustainability reports from emerging market countries are only available in the local languages, although some big companies, especially those with international operations or capital structure, would have full or partial reports in English, in addition to their local language. To overcome this issue, additional algorithms were developed to automatically identify and translate non-English content.

There are two main ways for this task. One is simply to translate the source into English using an automated translator, which implies in adding cost and complexity to the process, while also increasing the risk of imperfect interpretation by the AI algorithm, due to

translation imperfections and challenges in making sense of text even in the original language. For example, the following Portuguese phrases have very different meanings, although using almost the same words: (i) *No próximo ano, a indústria da moda promete mais transparência, aumentando o uso de rendas em suas coleções* and (ii) *No próximo ano, a indústria da moda terá mais renda e aumentará o investimento em transparência*. The first one tells that *next year the fashion industry will bring more transparent clothing, increasing the use of laces*, whereas the second tells that *next year the fashion industry will have more income and invest in transparency*. The reason is that some words are context-specific and may have different meanings in Portuguese, such as *renda* that can either mean *income* or *lace* and *transparência* (*transparency)* which, as in English, is an optical property applied by analogy in the management and sustainability context, with a specific meaning.

The second alternative is to teach the machine to interpret sustainability content in a given language. One way to accomplish that is to use deep learning algorithms with training sets in the local language to train the machine to derive meaning out of texts. This is not a trivial task, as it could take large input samples (of tagged text, for example) in the local language to get acceptable interpretation results from algorithms. Nevertheless, this task could be accomplished with a split-apply-combine strategy, which breaks the task into small pieces that together with crowdsourcing efforts can be accomplished at an accelerated rate (Salganik, 2018). No papers were found on text-mining sustainability reports in Portuguese.

Also in this alternative stills the issue regarding the specificity of terms in the sustainability field that affects the interpretation of machines, as exemplified above. This implies the need to develop a Portuguese sustainability lexicon to give additional context during the text mining process. For such purpose, a first necessary step would be the creation of an initial specific Portuguese sustainability lexicon to support training algorithms within a specific language field. Lastly, idiomatic idiosyncrasies and ambiguities are also another common issue in language processing.

Current use of BDAI by ESG information suppliers

Complementing group discussions in the roundtables, participating ESG information suppliers (ESG-IS) were interviewed in order to assess the current use of BDAI in this context, as well as to compare different methodologies presented.

In general, ESG-IS combine external sources of information (media, stakeholders, and public sources external to a company) with information disclosed by the companies themselves (Sustainability or CSR reports, company website, company codes of conduct and policies, management systems and certifications etc.). In some cases, external sources are compared with the ones provided by the firm, so that the ESG-IS can improve accuracy and detect greenwashing practices.

The interviews also revealed that ESG-IS are screening large amounts of publicly disclosed data on companies on monthly, weekly, and/or daily basis. Each ESG-IS has a unique combination of quantitative versus qualitative sources, and different ways regarding where to automate the analysis versus where to still rely on human analysts.

Although there are differences among ESG-IS data collection and processing methodologies, because they rely on BDAI technologies, they tend to face similar opportunities and obstacles. The main differences between them are best seen in the final product itself, that is, the product's features, to what type of client they are targeted and to what purpose are they useful. There are a variety of different approaches offered by current ESG-IS using BDAI, but in general, they tend to vary within a spectrum – and sometimes presented as a mix of both approaches – ranging from compliance, to CSR/sustainability management, to investment analysis tools.

BDAI presents substantial gains in terms of increasing the volume and quality of ESG information collected. However, as previously discussed, the complexity inherent in sustainability - which is properly addressed by Agenda 2030 and the SDGs - is a challenge for the effective incorporation of SDGs into BDAI tools for CSA.

There are already some efforts of ESG-IS to reconcile ESG information with the SDGs framework by mapping their ESG framework of choice to the 17 goals. For example, SASB and GRI have both mapped their respective themes to the SDGs (GRI & UN Global Compact, 2018; Eccles & Consolandi, 2018). However, there still a lack of research investigating if mapping ESG frameworks to the SDGs is a sufficient or appropriate approach.

It is clear that the lack of a robust framework linking the SDGs to current ESG frameworks limits the effectiveness of BDAI tools regarding an up-to-date assessment of companies in regards to sustainable development. This also hampers the potential of harmonization and improved comparability that a common framework could provide, which would strongly contribute to the usability of CSA tools.

CONCLUSION

The aim of this article was to understand the state of the art about ways of assess firm's sustainability performance using BDAI and also incorporating the more recent framework brought by Agenda 2030 and its SDGs.

Agenda 2030, especially SDG 12, brought to a new level of legitimacy the call for a broad, society-wide effort to build a sustainable and fair system of consumption and production. It draws attention to the urgency to reduce current its negative impacts on society and the environment and indicate that private companies have a special role in facing the challenge, commensurate to their resources and power of influence. Visionary investors, managers and entrepreneurs have been steadily improving their efforts to include environmental, social and governance aspects into their intelligence and decision-making systems.

In the past two decades, the demand for better disclosure from companies on their sustainability practices has increased, which led to the emergence of a new market for ESG information and an increase in the demand for sustainable investment. Despite the fact that there has been some efforts to guide sustainability disclosure, the inherent complexity of evaluating sustainability performance limits the objectivity, comparability and credibility of traditional Corporate Sustainability Assessment (CSA) and transparency tools, that relies mainly on yearly reports and limited media information. This limits their usability and may be hampering the acceleration of sustainable investment and management.

Despite having been originally created for governments, the SDGs present a potential opportunity to advance the harmonization of ESG frameworks and other CSA tools, providing a sound and legitimate common ground - a set of issues, objectives, goals, principles and guidelines - to which every methodology and tool should relate. A movement in this direction could lead to a solution for the above mentioned comparability conundrum, avoiding the risks and shortages of rigid standardization as well as of too loose guidelines.

Combined with that, BDAI seems to have the potential of radical improvements regarding the challenges of objectivity and credibility of information on firm's sustainability performance. By enabling the massive and uniform analysis of sources far beyond reports prepared by companies, it can prevent greenwashing and circumvent problems related to reporting voluntariness and biases. Algorithms have the ability to relentlessly search the internet for data from different sources and then, for instance, process the information

gathered in order to verify the congruence between the company's discourse and the perception of different stakeholders.

Although there have been progress in using SDGs and BDAI to advance the tools for CSA - such as mapping existing ESG frameworks to the SDGs or yielding relevant ESG data that can be converted into evidences with potential use by third party assessment and by the company itself - there are still many obstacles to be addressed.

For instance, as firms announce moving away from business as usual and proclaim their commitment to the SDGs, new questions arise, such as what exactly "contributing to Agenda 2030" means for a company? Or what "aligning a company strategy with the SDGs" actually is? The answers to questions as such are essential for the identification of companies more fit to the new business environment. This matters for companies managing sustainability performance and risks, and for investors who want to direct their resources towards more sustainable results.

In the BDAI front, a major issue relates to the lack of clarity regarding the comparability of methodologies that, in the absence of common references for effective harmonization, take very different approaches and may bring incongruent outputs. When it comes to Emerging Markets context, language is an extra barrier: the lack of sustainability *lexica* in local languages make NLP analysis unfeasible and worsens data quality, since translation demand additional resources and is not always accurate enough.

Another important consideration is regarding the access to information extracted with BDAI: although much of the available data comes from public sources, the ability to process and convert them into insights is costly, and there is a lack of publicly available tools to process the large volume of unstructured data. This may worsen knowledge asymmetries and aggravate power imbalancement.

For now, BDAI is likely to be more easily and effectively used as a means of providing better answers to current questions, rather than creating completely new parameters for a company's sustainability assessment. Nevertheless, the combination of BDAI with the common referential provided by Agenda 2030 and its SDGs seems to have the potential to overcome structural and persisting challenges of CSA and transparency.

BDAI tools that take into account Agenda 2030 are in their initial stages and will require time, resources and additional research to evolve into maturity.

The findings and conclusions presented in this paper are not decisive and should be taken as initial contributions for further studies in a new stream within the CSA research field. By sharing our insights into the present and glimpses of the future we hope to motivate researchers to advance towards what seems to us a promising way to face long standing challenges and build much needed new knowledge.

REFERENCES

Amel-Zadeh, A. (2017). Making the Match: The Market's Evolving View of the Materiality of Nonfinancial Information (p. 28). BNY Mellon. p. 18. Retrieved from http://www.ibri.com.br/Upload/Arquivos/novidades/3960_BNY-Mellon-Making-The-Match.pdf

Antolín-López, R., Delgado-Ceballos, J., & Montiel, I. (2016). Deconstructing corporate sustainability: a comparison of different stakeholder metrics. Journal of Cleaner Production, 136, 5–17. https://doi.org/10.1016/j.jclepro.2016.01.111

Arvidsson, S. (2014). Corporate social responsibility and stock market actors: A comprehensive study. Social Responsibility Journal, 10(2), 210–225. https://doi.org/10.1108/SRJ-08-2012-0099

Belinky, A. (2017). Da empresa cowboy à astronauta. *GV-executivo*, 16(5), 18-21.

BlackRock. (2016). EXPLORING ESG: A Practitioner's Perspective (VIEWPOINT No. GOV-0093) (p. 14). BlackRock. Retrieved from https://www.blackrock.com/investing/literature/whitepaper/viewpoint-exploring-esg-a-practitioners-perspective-june-2016.pdf

Borne, K. (2014). Top 10 Big Data Challenges – A Serious Look at 10 Big Data V's. MapR https://mapr.com/blog/top-10-big-data-challenges-serious-look-10-big-data-vs/

Bradford, M., Earp, J. B., Showalter, D. S., & Williams, P. F. (2017). Corporate sustainability reporting and stakeholder concerns: Is there a disconnect? Accounting Horizons, 31(1), 83–102. https://doi.org/10.2308/acch-51639

Brown, HS; Jong, Martin; Lessidrenska, T. (2009). The rise of the Global Reporting Initiative: A case of institutional entrepreneurship. Environmental Politics, 18(2), 182–200. https://doi.org/10.1080/09644010802682551

Castro, D. W., Souza, E., Vitório, D., Santos, D., & Oliveira, A. L. I. (2017). Smoothed n-gram based models for tweet language identification: A case study of the Brazilian and European Portuguese national varieties. Applied Soft Computing, 61, 1160–1172. https://doi.org/10.1016/j.asoc.2017.05.065

Eccles, R. G., & Consolandi, C. (2018, February 15). Supporting Sustainable Development Goals Is Easier Than You Might Think. Retrieved September 25, 2018, from https://sloanreview.mit.edu/article/supporting-sustainable-development-goals-is-easier-than-you-might-think/

Francisco, E. de R. (2015). A tendência do big data. *GV-executivo*, 14(2).

Filho, B., & Paulo, P. (2017). Aspect extraction in sentiment analysis for portuguese language (text). Universidade de São Paulo. https://doi.org/10.11606/T.55.2017.tde-05122017-140435

GRI, UN Global Compact and World Business Council for Sustainable Development (2015), "SDG compass: a guide for business action to advance the sustainable development goals", available at: https://sdgcompass.wordpress.com/ (Accessed 24 September 2018).

The International Integrated Reporting Council. (2013). The international Integrated Reporting framework. The International <IR> Framework, 37. https://doi.org/10.2139/ssrn.2378899

Irvine, M. (2016). The Sustainable Development Goals and Business, 5(2), 43. Retrieved from https://www.dnvgl.com/feature/sdgs-business-action.html

Jelodar, H., Wang, Y., Yuan, C., & Feng, X. (2017). Latent Dirichlet Allocation (LDA) and Topic modeling: models, applications, a survey. ArXiv:1711.04305 [Cs]. Retrieved from http://arxiv.org/abs/1711.04305

Jelveh, Z., Kogut, B., & Naidu, S. (2014). Detecting Latent Ideology in Expert Text: Evidence From Academic Papers in Economics. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP) (pp. 1804–1809). Doha, Qatar: Association for Computational Linguistics. Retrieved from http://www.aclweb.org/anthology/D14-1191

International Organization for Standardization. (2010). *Guidance on social responsibility* (ISO No. 26000:2010).

International Organization for Standardization. (2017). *Sustainable procurement - guidance* (ISO No. 20400:2017).

Liang, H., & Renneboog, L. (2017). On the Foundations of Corporate Social Responsibility: On the Foundations of Corporate Social Responsibility. The Journal of Finance, 72(2), 854–855. https://doi.org/10.1111/jofi.12487

Lozano, R., Huisingh, D., 2011. Inter-linking issues and dimensions in sustainability reporting. *J. Clean. Prod. 19*, 99e107.

SASB. (2018, August). Legal FAQs. Retrieved September 25, 2018, from https://www.sasb.org/approach/legal-faqs/

Mahadevan, V., Ng, A., & World Scientific (Firm). (2009). Proceedings of the 2009 International Conference on Software Technology and Engineering, Chennai, India, 24-26 July 2009. Singapore; Hackensack, N.J.: World Scientific Pub. Co. Retrieved from http://public.eblib.com/choice/publicfullrecord.aspx?p=1680998

Morsing, M., Schultz, M., & Nielsen, K. U. (2008). The "Catch 22" of communicating CSR: Findings from a Danish study. Journal of Marketing Communications, 14(2), 97–111. http://doi.org/10.1080/13527260701856608

Nielsen, A. E., & Thomsen, C. (2007). Reporting CSR - What and how to say it? Corporate Communications, 12(1), 25–40. https://doi.org/10.1108/13563280710723732

Porter, M. E., & Kramer, M. R. (2007). Strategy and Society: The Link Between Competitive Advantage and Corporate Social Responsibility. Harvard business review. 84. 78-92, 163.

Russo, A., & Perrini, F. (2010). Investigating stakeholder theory and social capital: CSR in large firms and SMEs. Journal of Business Ethics, 91(2), 207–221. https://doi.org/10.1007/s10551-009-0079-z

Salganik, M. J. (2018). Bit by bit: social research in the digital age. Princeton: Princeton University Press, p. 231-247.

Saraiva, M. (2017). Sustainability trends in Emerging Market Firms (p. 38). New York City: Columbia Center on Sustainable Investment (Unpublished).

Saraiva, M., Grajales, A. L., Macia, F., Goenaga, A., & Rosner, S. (2017). ALASKA RETIREMENT MANAGEMENT BOARD ESG Strategy: Ensuring Long-Term Returns for Alaskans (Final project) (p. 17). New York City: Columbia University.

Searcy, C., & Buslovich, R. (2014). Corporate Perspectives on the Development and Use of Sustainability Reports. Journal of Business Ethics, 121(2), 161. https://doi.org/10.1007/s10551-013-1701-7

Tschopp, D., & Huefner, R. J. (2015). Comparing the Evolution of CSR Reporting to that of Financial Reporting. Journal of Business Ethics, 127(3), 568-569. https://doi.org/10.1007/s10551-014-2054-6

US|SIF. (2016). 2016 Report on US Sustainable, Responsible and Impact Investing Trends (Trends Report) (pp. 12–13). The Forum for Sustainable and Responsible Investment. Retrieved from https://www.ussif.org/files/SIF_Trends_16_Executive_Summary(1).pdf

United Nations General Assembly. (2015). Transforming our world: The 2030 agenda for sustainable development.

Vendramini, A., Belinky, A. (2017). A nova fonte de valor econômico. GV-executivo, 16(5), 27-31.

Zadek, S. ; M., Mira. (2003). Redefining materiality: practice and public policy for effective corporate reporting. London: AccountAbility.

[Table 1]

Table 1: Examples of ESG information suppliers (ESG-IS).

Target market **Data Provider**	**Target clients**	**Market Coverage (#of firms)**
General market		
Bloomberg	Investors & Companies	>10,000
FTSE Russell	Investors	>4,000
MSCI	Investors	>6,000
Thomson Reuters	Investors	>6,000
ESG Exclusive		
Arabesque	Investors	>4,000
Covalence	Investors & Companies	>3,400
CSRHub	Companies	>17,000
Ethos	Investors	>1,650
Inrate	Investors	>2,600
Oekom Research	Investors	>3,500
RobecoSAM	Investors & Companies	>2,400
Sustainalytics	Investors	>6,500
VigeoEIRIS	Investors & Companies	>3,200
Specialized		
CDP	Investors	>2,000

Source: Douglas, Van Holt & Whelan (2017)

[Table 2]

Table 2 - Alaska Retirement Management Board Top Large Cap equities - Comparison of ESG Scores (2017)

Company	Sector	ASSET4	MSCI
JOHNSON & JOHNSON	Pharmaceuticals	95.39	BBB

DUKE ENERGY CORP NEW	Utilities	94.97	BBB
CVS HEALTH CORP	Retail - Food & Staples	94.2	B
KROGER CO	Retail - Food & Staples	94.16	BBB
TARGET CORP	Retail - Consumer Discretionary	94.09	BBB
HERSHEY CO	Food Products	93.95	BBB
AVALONBAY CMNTYS INC	Real Estate Management & Services	93.68	BB
DISNEY WALT CO	Media	93.44	BBB
EATON CORP PLC	Electrical Equipment	93.16	BBB
CONAGRA BRANDS INC	Food Products	92.91	BBB
VALERO ENERGY CORP NEW	Oil & Gas Refining, Marketing, Transportation & Storage	92.49	BBB
AT&T INC	Integrated Telecommunication Services	92.23	B
EXXON MOBIL CORP	Integrated Oil & Gas	90.22	BBB
PFIZER INC	Pharmaceuticals	89.47	CCC
HOME DEPOT INC	Retail - Consumer Discretionary	86.07	BBB
DARDEN RESTAURANTS INC	Restaurants	84.75	BBB
ZOETIS INC	Pharmaceuticals	75.57	AA

* MSCI scores range from CCC (lowest) to AAA (max). Asset4 scores range from 0 to 100.
Source: (Saraiva et al., 2017)