

Chapter 5

Synergizing Data Storytelling and AI: Enhancing Engagement in Complex Data Communication

Maria João Frade

Faculty of Economics, Universidade do Algarve, Portugal

Célia M. Q. Ramos

 <https://orcid.org/0000-0002-3413-4897>

CinTurs, ESGHT, Univeristy of Algarve, Portugal

ABSTRACT

Combining Data Storytelling concepts and Artificial Intelligence algorithms presents a powerful alliance for the analysis and communication of complex data, in a more intuitive environment and with higher levels of engagement. The main goal was to understand and present the potential of combining these two areas, where sentiment analysis was the method considered to represent Artificial Intelligence. The considered methodology consists, in a first phase of literary review, followed by sentiment analysis. The findings of this research culminate in the development of the definition of Data Storytelling, its potential, and its relationship with Artificial Intelligence systems, presenting a Dashboard Data Storytelling incorporated with AI.

INTRODUCTION

Throughout society, in general but especially in the business sector, it's easy to notice the shift in paradigm felt with galloping advances regarding data collection and processing and technology with access to AI, with its ability to influence, and

DOI: 10.4018/979-8-3373-6970-9.ch005

even make, important decisions. The development of capturing means and reading different types of data has generated such voluminous datasets (Big Data) that their treatment, extraction of insights and communication from them has become an almost impractical task without the intervention of Artificial Intelligence and advanced analysis and visualisation software. Consequently, resources such as Data Storytelling, among others, become fundamental to understanding this new reality, as well as providing a basis for its practice, according to the study *Big Data and AI-driven Product Design: A Survey* (Quan *et al.*, 2023):

“In recent years, big data and AI-driven product design have become a significant research hotspot. However, to the best of our knowledge, before this work, there has been almost no comprehensive summary of the applications of big data in product design. Therefore, a thorough literature review is necessary to provide theoretical foundations that can be utilised to develop scientific insights in this area.” (Quan *et al.*, 2023, p. 33).

The main challenge of appropriately analysing and transmitting the massive amount of data that has been intensifying is what Data Storytelling aims to solve. In addition to the problem of absorbing useful information, there is the added difficulty of communicating it. Feigenbaum and Alamalhodaie (2020) warn that:

“But while datasets and digital archives grow bigger and more open, information remains difficult to collect, complex to analyse, and challenging to communicate. As we’ve seen over the past two decades of this proliferation, more data does not necessarily lead to better data stories” (Feigenbaum & Alamalhodaie, 2020, p. 4).

This argument highlights the relevance of systems optimisation, and Storytelling associated with Data visualisation can be the solution to these demands. Data is one of the most precious business assets today, and companies that can convert their data into insights and those insights into knowledge will stand out, become smarter and outperform their competition. In a data-driven world, telling stories is a crucial aspect that will contribute to the success of organisations (Dykes, 2020).

Given and in addition to what was presented above, this study also aims to understand the impact these tools have on decision-making. The inclusion of Business Intelligence is fundamental. In this context, the main goals are:

- Familiarisation with the fundamental principles of Artificial Intelligence (AI) applied to data analysis, combined with Data Storytelling concepts.
- Present a case study related to the potential of combining these two areas and the creation of stories based on the insights obtained.

This study is divided into five sections, each one focusing on different aspects. Introduction establishes the context of the study, presenting the goals and the importance of the topic in focus, namely, the use of Data Storytelling and Artificial

Intelligence. Section two will focus on Artificial Intelligence, addressing its definitions, methods and the interconnection with Data Storytelling, and is fundamental to continuing the exploration of Artificial Intelligence. How can it improve data analysis and visualisation? The third section presents the methodology used to develop the study, followed in the text section by a detailed description of projects related to this topic. Finally, the last section presents the final considerations, the limitations of the study and presents future research paths.

DATA STORYTELLING AND ARTIFICIAL INTELLIGENCE

With the increase of Artificial Intelligence (AI) in data analysis, its approach in this study, and its intersection with Data Storytelling, are thus essential. AI has lately proven its potential and benefits to such an extent that AI-Powered Analytics products have been developed capable of analysing large amounts of data and extracting relevant insights in real time and interactively. The centre issue ends up being its communication, the way in which this information is transmitted and perceived. This communication must be streamlined through Data Storytelling, and AI must be able to use this technique to facilitate the transmission of the desired message.

Artificial Intelligence

Artificial Intelligence is a scientific area with quite some history that gained significant prominence in recent decades, and its rise is due to important milestones resulting from the work of several researchers. In the 1950s. British mathematician and computer scientist Alan Turing published an article (Computing Machinery and Intelligence) where he introduced the famous “Turing Test” as a way to determine the intelligence of a machine (Ergen, 2019: 6). During this period, the perceptron, a more primitive form of Artificial Neural Network (ANN), also emerged. “*The put forward perceptron by Professor Rosenblatt in 1956, before the basis of today’s artificial intelligence*” (Ergen, 2019, p.5). Hence, the consolidation of the term “Artificial Intelligence” - AI - as an area of study. In 1956, the term Artificial Intelligence was first coined by John McCarthy. Artificial Intelligence would be the science and engineering of making intelligent machines, especially intelligent computer programs (Morais *et al.*, 2020). Shortly thereafter, Samuel trademarked the term Machine Learning (ML) to describe systems capable of learning through the insertion and manipulation of data. “*One year later, after introducing the per-*

ceptron, American computer scientist Arthur Samuel coined the term 'machine learning.' (Ergen, 2019, p. 6).

In the 1960s, the concept of learning was expanded by stacking multiple perceptrons to create Deep Neural Networks (DNN) - [...] *resembling today's deep-learning architectures inspired from Neural Networks (NN) of the human brain.*" (Ergen, 2019, p. 6). However, interest in AI waned from the 1970s onwards, with few notable advances until the end of the decade.

In 1980, there was a resurgence of AI with Geoffrey Hinton, who developed the backpropagation algorithm, a fundamental technique for ANN training. At this time, other important techniques were also introduced, such as Convolutional Neural Networks (CNNs) for recognising presented images (Ergen, 2019).

The 90s brought some advances, such as Support Vector Machines (SVMs) for sentiment classification and Natural Language Processing (NLP). Recurrent Neural Networks (RNNs) and Long-Short Term Memory (LSTM) have also been introduced for advanced speech-to-text translation (Ergen, 2019).

In the early 2000s, Deep Learning (DL) was developed, an approach within AI's field that uses artificial neural networks (ANNs) with multiple layers that learn from data (Ergen, 2019: 6). Geoffrey Hinton was one of the pioneers introducing the concept of DL to the world, as well as demonstrating its abilities on several pattern recognition problems. To overcome some of its disadvantages and reduce training time, Transfer Learning (TL) was promoted, which consists of reusing knowledge from one problem to another related but different problem, using a pre-trained model (Ergen, 2019). *"In ML, you need to choose for yourself which features to include in the model. In order to overcome the drawback of insufficient data and reduce the training time, Transfer Learning (TL) is gaining popularity."* (Ergen, 2019, p. 7).

Nowadays, AI is increasingly present in our daily lives and plays a crucial role in many areas. Although there are concerns about its impact on jobs, politics, the economy, and the structure of society in general, it can bring many significant benefits if used responsibly and ethically. (Ergen, 2019).

"Expectations in AI will create more jobs than it will destroy. The AI hype will die eventually when AI becomes another thread as on the Internet, electricity, and transportation. Today, AI is promising, or in some cases threatens, changes in business and society at large. The industrial era let machines do the physical work; the information era enabled machines to do the computation and storage, now the AI era will let machines make the decisions" (Ergen, 2019, p. 7).

In more recent years, AI has spread and evolved due to its social and economic impact, especially due to the advances seen in DL. However, the complexity and opaque nature of DNNs make it difficult to understand the decision-making of these models, which limits their safe use. To solve this problem, eXplainable Artificial Intelligence (XAI) methods have been developed, but these have also been challenged.

This has driven Concept-based eXplainable Artificial Intelligence (C-XAI) methods, which are based on concepts and offer more understandable explanations, similar to human reasoning, making models more robust and effective (Poeta *et al.*, 2023).

In order to better comprehend the model, explanations based on C-XAI concepts offer an interesting alternative, as they provide a more holistic view of the inner workings of the model and the network. Poeta *et al.* (2023) state that: “*By explaining the model’s predictions in terms of human-understandable attributes or abstractions, concept-based explanations better resemble the way humans reason and explain.*” (Poeta *et al.*, 2023, p. 2). They also allow users to gain depth regarding the insights and their underlying reasoning, helping detect deviations from the model and improve their classification (Poeta *et al.*, 2023).

AI is often confused with automatization or digitalization, and is now everywhere, including most areas of science, but it is still far from having an appropriate universal definition, because on the one hand it belongs to the scientific field, on the other it has roots in biology, mathematics, psychology, philosophy and metaphysics, which makes its conceptualization a complex task. The definition of it as a concept has sparked discussion in the scientific community in recent years and has generated disagreements. “*A further inspection of the history shows that many have attempted to define AI. Every definition has been criticized and has failed to obtain consensus within each of these disciplines, let alone universal consensus.*” (Abbass, 2021, p. 95).

There is a need to ensure that the definition of AI can retain its true meaning and scope. The literature does not allow redundant definitions such as “AI is intelligence that is artificial”. “*We need to ensure that a definition is a “good” definition. The literature tells us that we should avoid self-referential statements; we should not offer a circular definition such as that “AI is intelligence that is artificial.”*” (Abbass, 2021, p. 94). It isn’t unreasonable that a definition of a term such as AI’s begins with exactly the two words that compose it: “intelligence” and “artificial”. One possible starting point can be beginning by defining these two words. For instance, intelligence as: cognition, learning, processing or gathering information; and artificial as: man-made, non-biological, imitating nature, unfaithful to nature, unfaithful to reality or manufactured (Abbass, 2021).

Thus, it is possible that AI is the automation of cognition, and the word “cognition” comes from the Latin word “cogito” (meaning “to know”), and that is not exclusive to humans. “*It could be the cognition of an animal or even the cognition of a different AI agent.*” (Abbass, 2021, p. 95). AI must be seen beyond its status as a technological product and more as a sociocognitive phenomenon. Although no theoretical definition of AI is yet error-free, sufficiently universal or concisely ambiguous, Abbass (2021), in his article *What is Artificial Intelligence?* presents the following definition:

“Artificial Intelligence is a social and cognitive phenomena that enable a machine to socially integrate with a society to perform competitive tasks requiring cognitive processes and communicate with other entities in society by exchanging messages with high information content and shorter representations.” (Abbass, 2021, p. 95)

Given the vast scope and capabilities of AI, the finding of a concise, clear and comprehensive concept that defines it is a difficult task. However, Morais *et al.* (2020) state the following:

“Uma Inteligência Artificial é uma estrutura composta e articulada por softwares e eventualmente, hardwares, cuja finalidade é auxiliar os seres humanos na tomada de decisão com base na associação de dados históricos e no reconhecimento de padrões”. (Morais *et al.*, 2020, p. 98).

Artificial Intelligence and Data Analysis

AI combined with Data Analysis is known as AI-powered Analytics. It is a proven but fairly recent technological advancement. Many companies have invested heavily in DL, a subcategory of ML (Machine Learning), which is in turn a subcategory of AI. DL and artificial neural networks enable image recognition, audio, NLP (Natural Language Processing), and other tools have been developed. Today, most modern Analytics solutions have recently started developing NLP that allows end users to use everyday language to ask questions about data. But it is undeniable that AI-powered Analytics will help organisations create and develop their most important strategies, and thus focus their business on the most important activities, instead of wasting valuable time with vast reports. *“AI and Machine Learning assist, augment, and amplify work. And when done right, these insights can empower everyone in the company, not just a few specialists.”* (Jyoti, 2021, p. 22).

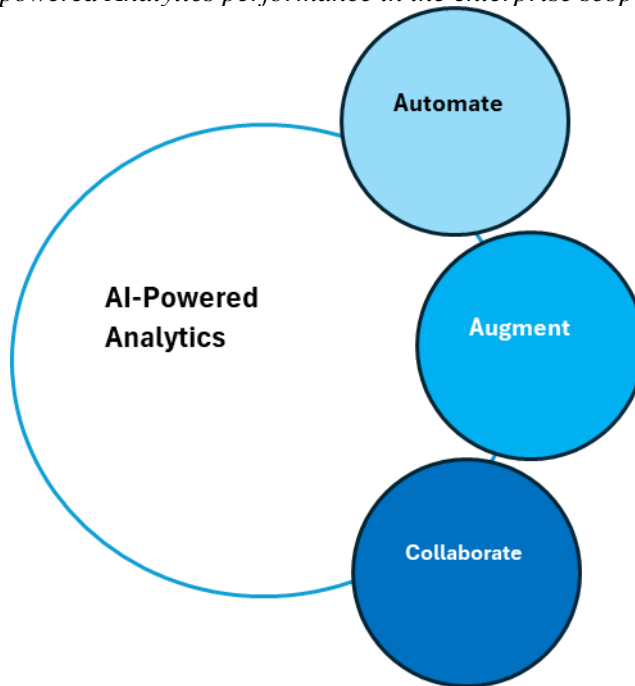
It is necessary to act concretely, that is, to collect and process all the data generated by consumers and use it so effectively that it exceeds everyone’s expectations. Analysis and insights are means to an end, not the end in themselves, and AI and ML allow, in addition to identifying the best action to take, also to automate some processes, which could translate into a benefit for everyone, companies and consumers (Jyoti, 2021).

Sometimes, developing and evaluating hypotheses and separating correlations from casualties to explain an unexpected anomaly or an upward trend can be a job that takes many hours and even days. AI-powered Analytics and ML techniques are able to classify billions of rows in seconds and show the indicators and main causes implicit in the data, guiding and increasing work efficiency, and giving greater consistency, accuracy and reliability to the results. These techniques are capable of presenting analyses and predictions with depth and detail unattainable by human

hand alone. They can, for instance, assume a hypothesis and abandon it if the data does not corroborate it, continuing in another more favourable direction until the best possible result is reached. These systems learn and become more intelligent as they receive more data and thus improve their ability to detect and correct anomalies over time (Jyoti, 2021).

AI-powered Analytics can help automate lifecycle tasks, augment human intelligence, or support human-to-human collaboration (Jyoti, 2021) as shown in Figure 1.

Figure 1. AI-powered Analytics performance in the enterprise scope



Source: Adapted from Jyoti (2021, p. 26).

Data analysis and AI algorithms allow us to discern what previously happened, project insights into the present, and make predictions about the future; that is, there are three types of analysis that can be performed (Liu, 2021):

Descriptive Analysis - The simplest type and usually the first step in the data analysis process, as it manages to consolidate data in a form that allows the generation of appropriate analyses and reports.

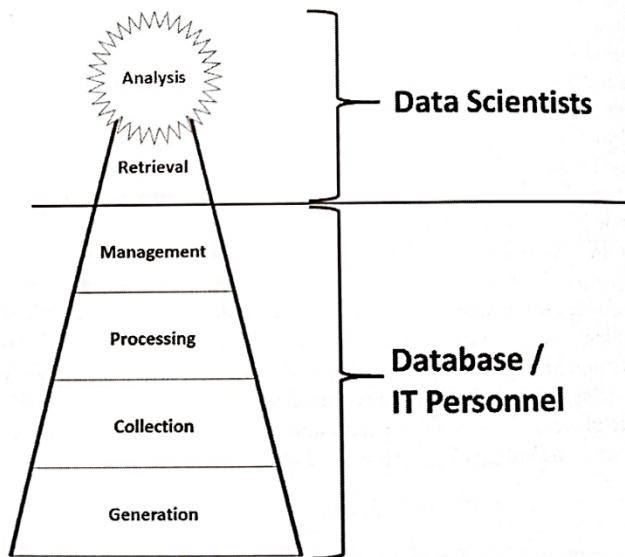
Predictive Analysis - Uses statistical models and ML to make predictions about future outcomes based on patterns found in existing data.

Prescriptive Analysis - Uses the results of previous analyses to create prescriptions or recommendations on the best actions to be taken to achieve the objective.

The essence of data analysis remains the same; what changes is the vast amount of data inside and outside organisations, as well as the enormous increase and development of computing power that can be used for its processing.

Data Science starts with data generation, which nowadays happens on a massive scale; everything is recorded, evaluated and stored. Data analysis and modelling are just some parts of the extensive data science cycle, as illustrated in Figure 2.

Figure 2. Data Science’s life cycle



Source: Nelson (2021, p. 10).

The comprehensive diversity of data types and their processing can bring as many benefits as challenges. Elements such as video, images or sounds must be indexed with keywords so that they can be processed, and a large amount of information will remain uncategorized. This can represent a major problem as long as organisations and companies do not realise the value of unstructured data (Nelson, 2021).

After obtaining and cleaning the data, the step that follows is to decide which data to use, understand what is needed for analysis and what may be missing, meaning, how one can find the data that reveals what one needs to know. This process known as Data Wrangling can be effectively assisted by AI, which can process all types of data types through the data lake (repository for storing Big Data, and structured and unstructured data). Nelson (2021) states:

“Data must be processed. Data Scientists must be able to trust the data provenance. [...] “Data wrangling” is the new data science buzzword. [...] However,

unstructured is proving difficult to manage. [...] AI can assist with this type of data by performing some of the wrangling by processing data through a “data lake”. (Nelson, 2021, p. 11).

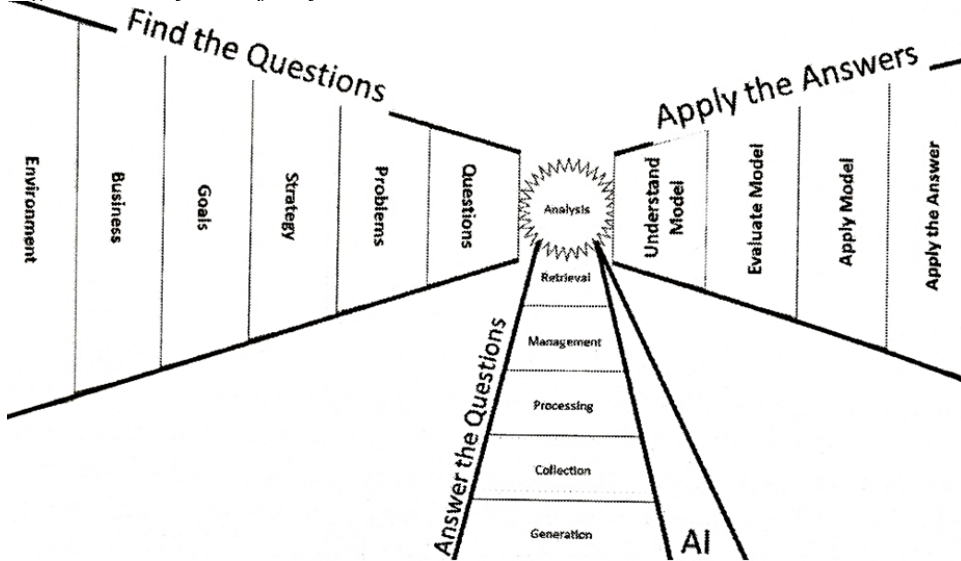
AI has the remarkable ability to examine and transform structured and unstructured data environments into relevant data, to identify important elements, categorise and purify these elements, and store them for later retrieval (Nelson, 2021, p. 7).

For these reasons, AI is essential to the future of data science and thus enters the field of Data Science Artificial Intelligence (DSAI) (Nelson, 2021). Therefore, the data science life cycle includes a very important aspect: understanding the business (Nelson, 2021).

Nelson (2021) argues that through the Analytics Life Cycle (ALC) all elements of analysis can be treated from start to finish in a business context, and divides ALC into 11 direct steps, shown in Figure 3, intuitive and useful that can be applied when you want to change something, especially if you want a change with some impact, and which are (Nelson, 2021, p. 8):

- 1- Comprehend the business environment.
- 2- Understand the business.
- 3- Understand the business's goal.
- 4- Comprehend the business's strategy.
- 5- Recognise and understand the business's struggles.
- 6- Develop questions.
- 7- Use data science life cycles.
- 8- Understand the analytic model through its visualisation.
- 9- Evaluate the analytic model.
- 10- Apply the model back to the questions (send unanswered questions by the model through the process).
- 11- Apply the answered question(s) to the business' problem(s).

Figure 3. Analytics Life Cycle



Source: Nelson (2021, p. 14).

These are some of the most popular approaches, in which analytics is seen as a “maturity journey” that progresses from the descriptive stage to the predictive stage, culminating in the perspective state. The synergies between AI and data analysis (AIA - Artificial Intelligence Analytics) are still a young science, but with a lot of potential, and professionals lack a lot of support. To address this shortcoming, Anderson and Zettlemeyer (2021) developed their framework, which they considered more suited to the real and current business world, shown in Figure 4.

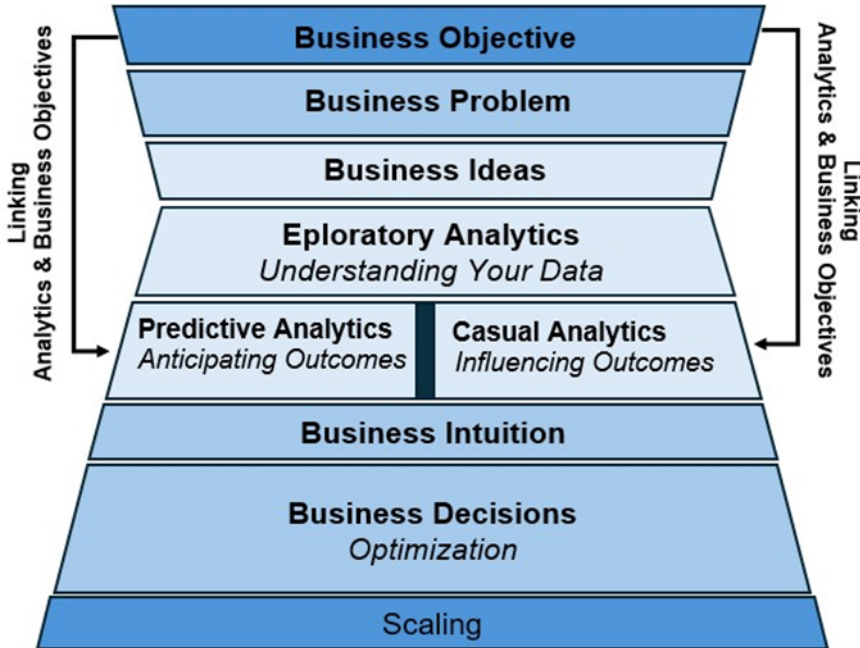
Anderson and Zettlemeyer’s (2021) framework focus on three specific types of analysis:

- **Exploratory Analysis:** Describes the data, explains how it was generated, and supports understanding the variability among the data.
- **Predictive Analysis:** Uses AIA models to anticipate future results.
- **Causal Analysis:** Uses AIA models to influence future outcomes.

The distinction between Predictive and Causal analysis is tenuous, but relevant. Firstly, because the business questions that can be addressed with Predictive analysis are different from those that can be addressed with Causal analysis. Secondly, the sources required to support a Predictive versus Causal analysis job can be quite different. Analysis is not the end goal; in order to determine the best action, it is

necessary to link the analysis with the business's goals (Anderson & Zettelmeyer, 2021).

Figure 4. Artificial Intelligence Framework for decision making



Source: Anderson and Zettelmeyer (2021, p. 27).

Artificial Intelligence and Data Storytelling

To relate AI and Storytelling, it is relevant to understand the importance and impact that communication has on humanity (Harari, 2011):

“That was the key success of sapiens. [...] In a conflict of hundreds, Neanderthals didn’t stand a chance. They could share information about the location of lions, but they probably couldn’t tell - and reframe - stories about tribal spirits. Without the ability to compose fiction, Neanderthals were unable to cooperate effectively in large numbers, just as they were unable to adapt their social behaviour to rapidly changing challenges. [...] The immense diversity of imagined realities that sapiens invented, as well as the resulting diversity of behaviour patterns, are the main components of what we call “cultures”. Once cultures emerge, they never stop changing and developing, and it is these unstoppable changes that we call “history”. (Harari, 2011, pp. 50-53).

As we continue today with the same eagerness, depending on, creating, searching for, and especially consuming stories. Imagining and communicating stories is inherent to the human condition. This reality is emphasised with the arrival of AI, which can learn and produce logical reasoning more effectively and quickly than the human brain, but cannot imagine and communicate stories with creativity, empathy and feeling, meaning, it is precisely the imaginary side of the human brain that distinguishes it as such. This conception must be taken into consideration and used to our advantage (Dykes, 2020).

Data visualisation and Data Storytelling focus on the perception and cognition of data, respectively. Therefore, data storytelling is expected to be applied in scenarios where analysis results need to capture their interest and trust. Compared to raw data, data stories are more easily memorable and cognitively more engaging. Additionally, they can align with naturally human characteristics and offer a higher level of experiential engagement. Throughout the narrative exploration of a story, there is continuous communication and interaction with the audience, feedback and revision of the plot, based on data (Daradkeh & Atalla, 2023). Thus, research on Data Storytelling takes on significant importance for data representation and the development of data products.

A data story is a type of data product or service that aims to meet the specific needs of a business by using data as raw material, applying data analysis and modelling methods to extract valuable insights from it, and presenting them in the form of a story to a determined target audience. The process of transforming data into a data story is called Data Storytelling. A data story is the end product of a Data Storytelling process, combining the objectivity of data and the subjectivity of the story. A data story starts from the needs of a targeted business, always relies on data as evidence, and involves sequential stages of analysis.

Currently, applying Data Storytelling can be organised into two main categories (Daradkeh & Atalla, 2023):

- **Direct data stories Generation:** this genre develops and delivers data products in story form to the audience, such as data journalism, dashboards, and academic presentations.

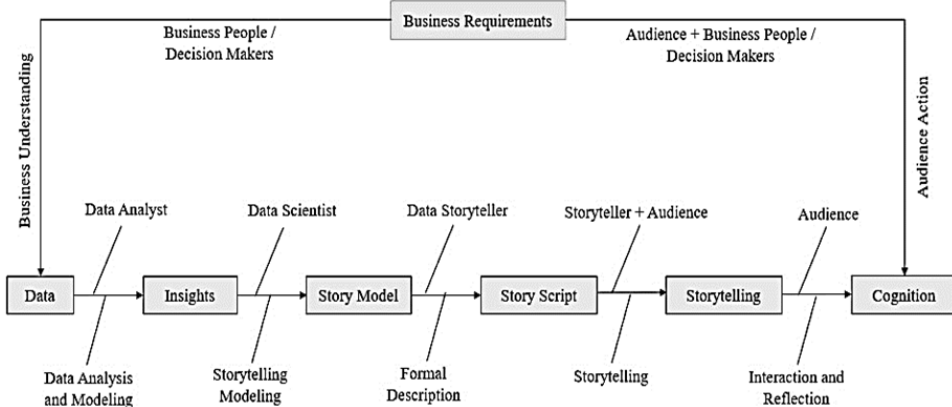
- **Data Storytelling Software Product Development:** This type provides and implements methods and techniques to support platforms or tools used to generate data stories. Currently, Data Storytelling platforms or tools can be subclassified into two categories: those that incorporate Data Storytelling capabilities into Data Visualisation or Business Intelligence, such as Tableau, TIBCO Spotfire, Flexdashboard, Qlik Sense, ShortHand and Microsoft Power BI; and those that are specifically developed for Data Storytelling, such as Banjo. Generally, Data visualisation software that supports Data Storytelling products, and the number of specialised tools designed specifically for Data Storytelling is relatively limited.

This last genre of Data Storytelling is where the core of this work lies. In this type of Data Storytelling - Development of Data Storytelling Software Products - is where the intersection with AI occurs, where one can observe the performance of the algorithm in the creation of narratives, and more specifically, in visual narratives. In other words, it falls in the first subcategory: Development of Data Storytelling Software Products that incorporate Data Storytelling capabilities into Data Visualisation or Business Intelligence tools. According to Daradkeh and Atalla (2023):

“Currently, data storytelling platforms or tools can be classified into two categories: those that incorporate data storytelling features into existing data visualization tools or business intelligence software, such as Tableau, D3.js, TIBCO Spotfire, Flexdashboard, Qlik Sense, ShortHand, and Microsoft Power BI; and those that are specifically developed for data storytelling, such as Banjo and Narratives for Tableau. Generally, visualisation software that supports data storytelling is the mainstream approach in the development of data storytelling software products, while the number of specialised tools specifically designed for data storytelling generation is relatively limited” (Daradkeh and Atalla, 2023, p. 551).

From a perspective of automatic generation and engineering refinement, the data story creation process can be decomposed into six main activities: understanding the data and the business, analysing data and insights, modelling the narrative, formalising the storytelling description, narration and interaction, and audience interaction and reflexive behaviour. According to Dradkeh and Attalla (2023), such as illustrated in Figure 5.

Figure 5. Main activity of Data Narratives

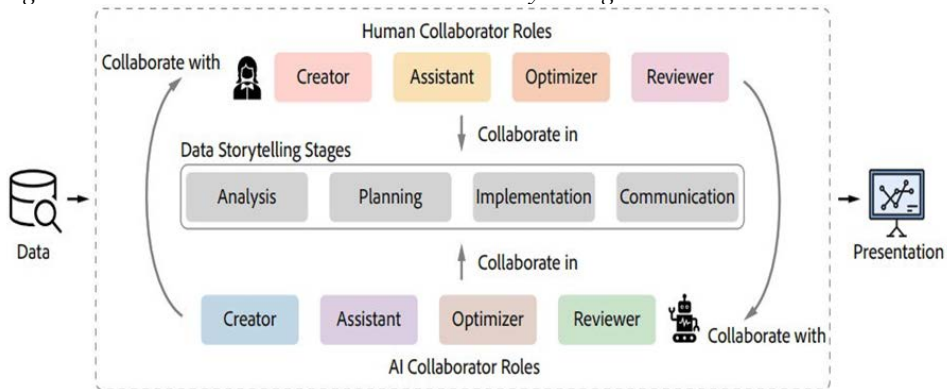


Source: Daradkeh and Atalla (2023, p. 552).

Crafting visually appealing data stories requires creativity, logic, sensitivity and other diverse skills, data analysis being one of them, allied to a keen awareness of the audience and context (Li *et al.*, 2024), however there remains a gap in literature on Data Storytelling tools, especially from a human-AI collaboration perspective (Li *et al.*, 2024).

On human-AI collaboration in existing Data Storytelling tools, Li *et al.* (2024) cite the importance of the different detailed stages: data analysis, data story planning, data story implementation, and its communication (Li *et al.*, 2024). And at each stage, collaboration patterns are examined through creators, reviewers, optimisers and assistants (Li *et al.*, 2024: 2)- Separating roles makes it easier to understand how humans and AI contribute to each stage of Data Storytelling. When combining these aspects, it is possible to comprehensively understand the design of human-AI collaboration. Figure 6 illustrates the framework of human-AI collaboration in Data Storytelling tools. It characterises collaborative data storytelling tools between humans and AI, based on the phases they encompass and the role they partake in these collaborations per phase.

Figure 6. Human-AI Collaborative Data Storytelling Tools



Source: Li *et al.*, (2024, p. 1).

The distinctive roles - creator, assistant, optimizer and reviewer - for a worthy understanding of this process, translates into the following matrix that defines the four types of human-AI collaborations (Li *et al.*, 2024), where the coding of the collaboration patterns between humans and AI in Data Storytelling is done by the author himself and which are represented in Figure 7.

Figure 7. Four types of Human-AI Collaborations

Role	Definition
Human-creator , AI-creator	The collaborator who finishes most of the work from scratch
Human-assistant , AI-assistant	The collaborator who works with creators together to reduce the workload or compensate for inability
Human-optimizer , AI-optimizer	The collaborator who automatically improves the entire or part of the data stories
Human-reviewer , AI-reviewer	The collaborator who assesses the content in data stories and provides assessment results or suggestions

Source: Li et al., (2024, p. 5).

Researchers have extensively investigated collaborative ways between humans and AI to automate Data Storytelling whilst considering human control. This aligns with the creative nature of Data Storytelling, where designing stories should be taken as a creative task and it is difficult to define a clear goal to be achieved automatically by AI (Li et al., 2024), since the creation of data stories still rely on the experience and judgement of humans, it is crucial to have human control throughout the Data Storytelling process, and simultaneously optimize the automation process.

In conclusion, there is interest and many advances regarding applying Data Storytelling in AI systems; however, this process still depends on the human touch to give it the genuineness that distinguishes it and makes it effective. Li et al. (2024) also propose two directions: exploring more design options in collaborative human-AI data storytelling tools and considering storytelling tools in conjunction with general visualisation creation and creativity support tools.

Data Storytelling and Artificial Intelligence for Data Analysis

The conjunction of the disciplines of Data Visualisation and Data Storytelling gives rise to the terminology of Visual Storytelling, as previously noted in this work. According to concepts stipulated in the study by Daradkeh and Atalla (2023, p. 552), and discussed previously in this study, Visual Storytelling is: “A primarily visual storytelling approach that emphasises the importance of visualisation techniques in data storytelling.”

Popular sayings such as “a picture is worth a thousand words” or “judging a book by its cover” are based on the common and widespread observation of human behaviour in society. If the story is the key element to captivate an audience, visualisation is no less important. And sometimes, more important than the plot of the story is the way it is presented. If the challenge in data analysis is - how to express an insight through numbers; in data visualisation - how to present data visually; in Data Storytelling - how to tell a story with data that expresses the desired insights; in Data Visualisation, the challenge will be - how to tell that data story visually.

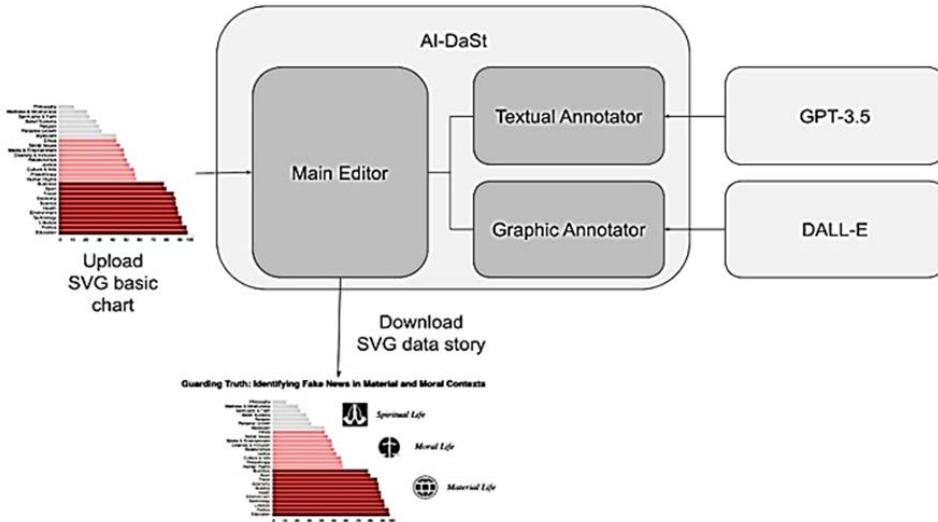
In the wide range of communications, there are different types and elements, and specifically, visual communication encompasses verbal communication (text) and non-verbal communication (images). Although visual communication of data always requires some textual communication, it should be strictly the minimum and necessary; graphics and images should prevail, highlighting their immediate and intuitive nature, which allows for faster, easier and clearer interpretation (Knafllic, 2015). Within the images, there are also animated images and videos that, like other visual elements, must be carefully inserted.

Generative AI, although recent, can be found in the task of automatically completing emails, for example, and therefore it is mostly used in co-writing and education. Lo Duca's (2023) study then focuses on how to assign a narrative to a data product in terms of annotations and introduces Generative AI to enrich data visualisation. *“Applying data storytelling to data visualisation means adding a narrative that better explains the visual and engages the audience. Generative AI can help transform data visuals into data stories”* (Lo Duca, 2023, p. 512).

Generative AI has opened new possibilities for the enhancement of data visualisation with narrative elements. Generative AI is a subfield of AI that uses Large Language Models (LLM) to generate new text based on a given prompt. It can be useful for different purposes, including text, audio and image generation (Lo Duca, 2023). There has been an effort made to identify the most common approaches to constructing stories in the media and news, and seven genres of visual narrative for journalistic stories have been recorded: magazine style, annotated graphic, partitioned poster, flowchart, comic book style, slideshow, and video (Lo Duca, 2023). And the semantics of the content of textual descriptions of the graphics is organised into two levels: enumerating the visualisation construction priorities, reporting statistical concepts and relationships, identifying perceptual and cognitive phenomena, and explaining domain-specific insights (Lo Duca, 2023).

AI Data Storytelling (AI-DaST) is a framework that allows for improving data visualisation and graphics through Generative AI. Figure 8 shows the AI-DaST architecture, which comprises three main elements: the Main Editor, the Textual Annotator Editor, and the Graphical Annotator.

Figure 8. AI Data Storytelling Architecture



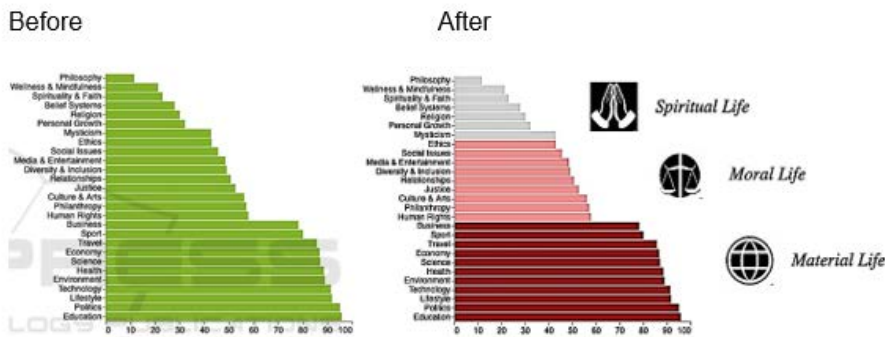
Source: Lo Duca (2023, p. 514).

The AI-DaST architecture considers four types of annotations (Lo Duca, 2023):

- Title - a concise sentence that summarises the main purpose of the chart, and the way it is written influences its interpretation.
- Comment - a brief comment providing additional information or context to the chart.
- Note - a brief written observation that gives additional information about a specific point in the chart.
- Image - a visual representation of the main subject of the graphic, an image that helps to engage the audience from an emotional point of view (Lo Duca, 2023: 514).

For a better understanding of these concepts, Figure 9 reveals the differences between graphs with and without AI-DaST. That is, the left side presents an initial graphical representation, and the right side presents the graph after using the Graphic Annotator.

Figure 9. Applying Data Storytelling in Graphics (Before and After)



Source: Adapted from Do Luca (2023, pp. 516 - 517).

In this example by Do Luca (2023), the automatic change in the way the graph is read after applying Data Storytelling is noticeable, as shown in Figure 9. Compared to the original graph (on the left), the graph that represents Data Storytelling techniques is much more appealing, enlightening and quickly understandable (on the right). The main differences reside in the added elements, but also in the way they are integrated and add perceptual features. Applying short, highlighted text that serves as a caption (Textual Annotator), of colours to create structured divisions and of icons as signage (Graphic Annotator), instinctively guides the eye towards faster, more informed and useful reading.

The improvement of Data Storytelling in AI systems is the concluding point of this section, which takes as an example and establishes a connection with another developed by Do Luca, who states that in the future, it may involve the additional refinement of the AI-DaST framework, addressing scalability challenges and exploring additional forms of applying AI in data storytelling.

METHODOLOGIES

The methodology used for this study has a first stage, where the research and consultation of relevant bibliographic references for the acquisition of knowledge about Data Storytelling and the interconnection with Artificial Intelligence will take place.

For this study, the following stages, mainly qualitative methods, will be used. In particular, the participant observation technique will be used as a reference to implement Data Storytelling and Visual strategies.

Finally, a methodology will be presented to combine Artificial Intelligence, through the consideration of Sentiment Analysis, with concepts of Data Storytelling.

Data Storytelling

The first stage consisted of defining Data Storytelling, based on the literature review. For this purpose, and after a set literature review, the following definition was considered: Data Storytelling is the process of transforming raw data into understandable and engaging narratives by combining data analysis with narrative and visual elements. Its purpose is to facilitate the understanding and retention of complex information, helping the decision-making process and engaging the public. For doing so, it uses graphics, charts and visual design to present insights in a clear and persuasive form, making it an essential tool in business, journalism, education and public health.

Sentiment Analysis

In the second stage, the goal is to develop a project where an artificial intelligence algorithm will be applied to Sentiment Analysis on satisfaction comments about an event, as a demonstrative example of the potential of combining AI with Data Storytelling, an approach with ML that is a subcategory of AI, as a “proof of concept”.

Sentiment Analysis (SA) (also known as Opinion Mining, Opinion Extraction and Subjectivity Analysis) focuses on the extraction of people’s opinions, emotions, feelings and attitudes toward entities such as organisations, brands, personalities, services, products, issues, events or topics (Liu, 2021). With the growth of social media databases on the Internet, the amount of opinion data in digital form is as large and accessible as ever. Therefore, since the 2000s, SA has become one of the most interesting research topics in the field of Natural Language Processing (NLP). Due to its important contribution to decision-making, applying it has been spread to areas such as Marketing, Finance, Politics and Health, and many companies are currently dedicated to providing SA services. Several applied studies have demonstrated its useful contribution in predicting sales performance, election results and market trends (Liu, 2021).

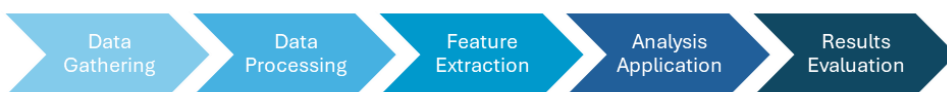
Sentiment Analysis can be considered and applied at three levels (Liu, 2021):

- **Document Level:** involves classifying an entire document as expressing a positive or negative sentiment. This approach is limited to documents that evaluate a single entity and is not suitable for texts that compare multiple entities.

- **Sentence Level:** Focuses on analysis at the level of each sentence, identifying whether it expresses positive, negative or neutral opinion, it is associated with the subjectivity classification that distinguishes factual sentences from subjective sentences. However, they still don't capture all the nuances.
- **Entity and Aspect Level:** Performs a more refined analysis by specifically identifying which aspects of an entity were appreciated and the ones that were criticised. Rather than analysing documents or sentences as a whole, this level looks directly at the sentiment associated with entities and their individual attributes. This level transforms unstructured text into structured data, facilitating detailed qualitative and quantitative analyses.

SA has several steps to be completed for its conclusion; these steps constitute the methodology for applying SA to comments. The author El-Masri *et al.* (2017) present a schematic methodology for the SA process, as shown below in Figure 10.

Figure 10. Sentiment Analysis Methodological Process



Source: El-Masri *et al.*, (2017, p. 3).

The different steps that are part of the SA methodology are as follows (El-Masri *et al.*, 2017):

- **Data Gathering:** Opinions can be collected from a variety of sources, the most common currently being social networks such as Facebook and Twitter, which are multi-domain and contain opinions on different topics. It is also possible to collect data from specialised sites, such as TripAdvisor for tourism reviews, IMDB for movie reviews and Amazon for product reviews.
- **Data Processing:** Processing means the process of cleaning data in order to remove unwanted elements, thereby increasing the accuracy of results and reducing the chance of errors. However, excessive use of these techniques can lead to the loss of important information. Processing techniques include:
 - Tokenise text: Split and categorise a text, words, sentences or characters.
 - Remove stopwords: Words that are irrelevant to sentiment analysis are eliminated, which improves effectiveness.

- Remove or identify punctuation: Punctuation marks are generally removed, but some, such as the exclamation point, can indicate emotional polarity.
- Convert text to lower or upper case: Capital letters can suggest strong emotion.
- Stemming the word: Reducing words to their basic form, a more complex process in certain languages such as Arabic and Chinese.
- Check the Spelling: Correct common typos, especially in social media data.
- Letter replacement: substitutions can be made to improve the accuracy of the data analysis. For instance, Twitter data requires additional techniques, such as removing hashtags, ERULs, and user mentions, as well as dealing with emojis and specific characters. Dialect manipulation is also important. In some cases, “unique” tokens are added for rare words in short tweets. And all these techniques help transform unstructured text into structured data, facilitating qualitative and quantitative analysis. Some of the text enrichment techniques in NPL are part-of-speech tagging (identifying grammatical functions of tokens), and syntactic analysis (building the syntactic structure of sentences) (Liu, 2021).
- **Feature Extraction**: Feature extraction allows for more accurate sentiment analysis and detailed summarisation of results. The most common features in SA are n-grams and POS (Part of Speech) tagging. Due to the complexity of the Arabic language, some studies have explored other specific linguistic features. The extracted characteristics can be of various types: stylistic, syntactic, semantic, lexical, and other specific needs.
- **Analysis Application**: There are three main methods of SA: Machine Learning, Lexicon-based, and hybrid or combined methods. In SA, the concepts of polarity (which can be positive, negative or neutral), subjectivity (which refers to feelings and opinions) and objectivity (referring to information, concrete data and facts) stand out (Costa, 2018).
- **Results Evaluation**: The evaluation of the results must be done taking into account the objective/problem initially defined, and possible limitations such as: the correct recognition of dialects, expressions and figures of speech (metaphor, sarcasm, irony, etc.); translation or interpretations errors; inability to compare opinions with each other; the lack of literary enrichment and colloquial language; and the appropriate classification of feelings and subjectivities. (El- Masri *et al.*, 2017).

The extractions of information for effective SA depend on the information gathered for it, since in this study, it was not possible to use company data; only the methodology to be considered for future studies is presented.

Based on the notions of SA, this point aims to highlight the importance and explore the methodology of applying Data Storytelling in the presentations of results, which integrate results from sentiment analysis. And to these ends, the proposed framework is as follows:

1. Collect data from social networks which represent consumer opinions, corresponding to the Data Gathering Phase.
2. Clean and process the collected data so that it can be analysed, corresponding to the Data Processing phase. This stage includes the following steps (Liu, 2021):
 - a) Tokenisation: Break text into meaningful tokens.
 - b) Stop words removal: Eliminate words that do not add value.
 - c) Normalisation: Standardise tokens and convert uppercase to lowercase.
 - d) Stemming and lemmatisation: Reducing words to their primary form.
 - e) Spell correction: Correct typing errors.
3. Organise/Identify what to analyse in terms of characteristics or aspects to be investigated, corresponding to the Feature Extraction phase.
4. Applying the Sentiment Analysis Algorithm.
5. Assessment of SA results.
6. Present the SA results on dashboards, according to Data Storytelling concepts presented above.
7. Integration of other data in the company, such as financial data and development of dashboards that allow the analysis and comparison of SA results with financial results, detecting possible patterns in them.

SA is thus defined as something complex, something where it is of importance to structure the problem to objectively identify and understand effects and causes. This structured definition makes it easier to solve the challenges associated with unstructured natural language text and to analyse a wide range of opinions, reducing their subjectivity and combining them with emotions.

APPLYING AI WITH DATA STORYTELLING

Opinions are crucial for the decision-making process and influence our behaviours and choices, and SA studies how feelings, evaluations, attitudes and emotions shape our perceptions. The task of finding and analysing these opinions continues to be challenging due to the large volume and diversity of the available data, and

therefore, investment and research in automated SA systems become increasingly essential (Liu, 2012).

Towards this end, we proceeded to apply a satisfaction survey to a considered event, where the concepts of Data Storytelling were applied, combined with AI methods. This process took into consideration the Sentiment Analysis methodology proposed by El-Masri *et al.* (2017) described above, which includes the following steps:

1. Collect

Data was collected through a satisfaction survey carried out with event participants. As they contain responses with text and are subject to SA, the following questions were considered for the purpose of this Sentiment Analysis:

Question 2: What were the main reasons for your rating above?

Question 3: Were there any specific sections or topics you found particularly valuable? Please mention them.

Question 4: Were there any specific sections or topics you found lacking in quality or relevance. Please mention them.

2. Clean

This particular task began with cleaning up the text. Special characters, emojis, irrelevant words, and extra blanks were removed, as well as some spelling corrections that were made, thus preparing the text for the next stage.

3. Process

When processing this data, normalisation was carried out to convert all words to lowercase. It was also necessary to remove stopwords.

4. Organise

Since the only goal is to present a “proof of concept”, only a small set of responses (16 in total) were considered, where the aspects analysed were not identified, leading to only the application of Sentiment Analysis at the global comment to be considered.

5. Applying the Algorithm

The algorithm was applied, tested and trained to create an analysis model. In this context, the chosen algorithm was VADER, given that it is considered the most suitable for analysing comments made on social networks. – “*Interestingly, the VAD-*

ER lexicon performs exceptionally well in the social media domain. [...] it is more sensitive to sentiment expressions in social media contexts while also generalizing more favourably to other domains.” (Hutto & Gilbert, 2014).

6. Results Analysis

In the analysis of the results of this SA, the examination is carried out question by question (2,3,4), highlighting the best and worst results, and the average. Word clouds were also generated based on the obtained responses, allowing the visualisation of the most used words and a clearer perception of the expressed sentiment.

RESULTS

Data collection was carried out through a digital satisfaction survey aimed at event participants. In total, 16 surveys were answered. Based on the responses obtained, analyses were carried out that correspond to the questions in the survey, to understand public satisfaction and identify aspects that can be improved in the organisation of a future event of a similar nature.

Question 2 - What were the main reasons for your ranking above?

In this question, the answer with the highest sentiment was *“Very enjoyable conference with opportunities to make it even better!”*, with a value of 4.5. The response with the lowest sentiment was *“Nice programme, but announcements not very clear. Also missing a coffee/tea break in the afternoon and I think there should have been a bar/drinks place open at the venue, even if we would have to pay.”*, with a value of 0.7. The average sentiment recorded was 2.9, and Figure 11 represents the word cloud generated from the collected responses, showing *“Good”*, *“Conference”*, *“Organization”* and *“Speaker”* as the main influencers of the overall evaluation of the event.

Question 4 - Were there any specific sections or topics you found lacking in quality or relevance. Please mention them.

Please note that in the following answers, there are concrete names, such as speaker names, which have been replaced by “###” to protect their identity and comply with current data protection policies.

The answer with the highest sentiment for this question was “### - is it for normal people like me? It seems really high-end and the speaker kep saying that we could contribute and this project is for us - honestly couldn't see that. In what way can one participate and how can we benefit from it.”, with a score of 4.5, the one with the lowest sentiment was “Yes, ### and ### presentations were bad.” Figure 13 shows the word cloud based on the text resulting from the provided answers, which reveals “Presentations”, “No”, “All” and “Speaker” as the most representative terms.

Figure 13. Word cloud generated through responses to question 4



Source: Own elaboration on Orange Data Mining.

In this analysis, the Orange software and the VADER algorithm were used, as already mentioned. This algorithm performs SA based on rules and a lexical dictionary: “[...] then combine these lexical features with consideration for five general rules that embody grammatical and syntactical conventions for expressing and emphasising sentiment intensity. Interestingly, using our parsimonious rule-based model to assess the sentiment.” (Hutto & Gilbert, 2014). VADER only works with text in English since all the answers were given in English; in this case, there is no need for translation. To start this process, it is necessary to load the corpus file so that it can be linked to the SA, and by calculating the negative, neutral and positive

scores, the column with the value resulting from the combination of all these scores (sentiment) is generated.

Sentiment is expressed from -1 to 1, where a value of -1 expresses the most negative sentiment and +1 expresses the most positive sentiment. To facilitate reading and analysis of the data, as well as comparison with other scales, the sentiment value was converted to a scale from 0 to 5 (where a value less than 2.5 represents negative sentiment and a value greater than 2.5 represents a positive sentiment). The conversion was performed using the following equation (Chen, 2020).

Based on the analysis of the results presented, it can be concluded that the overall balance is favourable. The majority of responses were indicative of people being pleased with the experience and its attributes. In general, the graphs represent positive values, revealing negative responses on a smaller scale. The level of “Very Satisfied” is reached in almost every graph, and although evident and not residual, it allows us to affirm a satisfactory evaluation of the event by the participants surveyed.

Dashboard Results

To present these results in a more readable and attractive way, a graphical representation was used, and the software chosen was Microsoft Power BI.

The goal here is to create a dashboard with the SA results, based on the responses given to the satisfaction survey. The aim is to apply the acquired knowledge coherently, develop the skills of crossing the studied tools, and highlight their usefulness and importance.

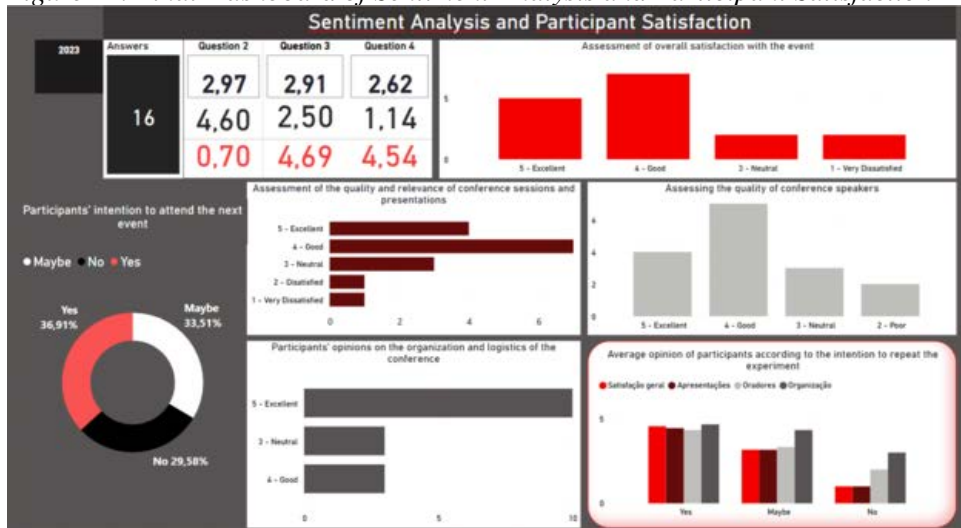
Dashboards are essential for the synthesis of large volumes of data into understandable visualisations, facilitating analysis and strategic decision-making, and for this reason, they play a fundamental role in competitive management. Through Data Storytelling, dashboards transform large amounts of complex data into clear and accessible visualisations, promoting faster, deeper and more enlightening assimilation of data sets. In turn, AI enhances not only the ability of dashboards to provide advanced information, such as predictive analysis and automated insights, which allow companies to respond quickly and conveniently to the challenges they face.

Figure 14 presents the dashboard built to communicate the results of the SA carried out on the responses from the event participants who took part in the satisfaction surveys, displaying Visual Storytelling techniques that apply in several aspects. Regarding text, titles were added and adapted for quick understanding of the graphs, with colour and font suited to the easy reading and general aesthetics of the panel, which were stylised for more intuitive communication. Colours were chosen according to their ability to facilitate highlighting and suggestion of automatic associations of the issues being analysed in the final graph that interconnects them, also highlighted. Dashboard data is linked in a way that allows cross-highlighting

between its elements, making evident the selected results in all graphs simultaneously, and allowing perception of the related components.

One of the Data Storytelling focuses of this dashboard lies in the positioning of the visual elements. In line with the common reading orientation, from left to right, it begins with the graph corresponding to the question about the intention to repeat the experience of the event again, with the other graphs on the right that analyse different isolated categories, and ending with the graph that interconnects them and shows their relationship, highlighted in colour in the lower left corner.

Figure 14. Final Dashboard of Sentiment Analysis and Participant Satisfaction



Source: Own elaboration on Microsoft Power BI.

This dashboard, based on the previously discussed Data Storytelling practices, allowed for adequate graphical visualisation of the results, as well as the relevance of applying AI tools to data processing. The combination of Data Storytelling with AI has proven to be an operational collusion for presenting complex insights in a concrete way, allowing for objective communication of results.

CONCLUSIONS

This work begins with a literary review on Data Storytelling, its definition and the combination of its application with AI tools. Starting by understanding what AI and Data Analysis are, what they are for, and how they work.

Data Storytelling refers to transforming data into engaging narratives, combining analysis and graphical visualisation to facilitate decision-making and understanding complex information. AI, in turn, is introduced to assist and improve the analysis and presentation of data, being applied, for instance, in Sentiment Analysis and in the creation of dashboards that allow for more flexible interpretation of Big Data.

AI can also be integrated into data analytics processes, creating accurate predictions about customer behaviour, providing businesses with a competitive advantage. In short, combining these tools can not only improve a company's analytical capabilities but also provide a more effective and accessible understanding and communication of the insights generated.

Nonetheless, the most striking limitation of this study was the lack of data to obtain more interesting results. The lack of access to some of the company's key data did not allow other types of information to be included in this study.

One of the topics worthy of deeper investigation would be the distinction, and respective systematisation, between "creation" and "extraction" of Data Stories. Data-driven stories can be drawn from a dataset, but they can also be adapted or fictionalised and still be based on factual data. Comprehending these differences and their methodology would also be important to decide when to use software with Data Storytelling resources (e.g. Microsoft Power BI), or one that is specific to carry out Data Storytelling (e.g. Banjo). Another aspect to be included would be the concepts associated with color and their effect on users' emotions.

Since these tools are advancing and seen as indispensable in the near future, it would be very important to develop a study aimed at Generation Z to examine their relationship with data visualisation.

A more detailed study on the influence, the application and the good practices of data visualisations with AI to transform complex information could also result in a valuable contribution, understanding what can enhance engagement, attention and empathy of the general public. As such, it is important to investigate phenomena like this, whether from an academic, artistic or social point of view.

The study's reliance on a small sample size (16 survey responses) significantly limits the generalizability of its findings. While the sentiment analysis provided preliminary insights, the narrow dataset risks overlooking nuanced patterns or biases inherent in larger, more diverse populations. A critical examination reveals that such constraints may skew results, particularly in sentiment analysis, where cultural, linguistic, and contextual variability play pivotal roles. Future research should prioritise

larger, stratified samples to enhance reliability and capture broader sentiment trends, ensuring findings are robust and actionable for real-world applications.

The methodology, though well-structured, could benefit from deeper exploration of human-AI collaboration, particularly in how biases in AI algorithms (e.g., VADER's English-language limitations) interact with human interpretation. Ethical considerations—such as data privacy, transparency in AI decision-making, and the risk of automating subjective judgments—warrant further discussion. For instance, while AI streamlines sentiment analysis, human oversight remains crucial to validating results, address ambiguities (e.g., sarcasm), and ensure ethical data usage. Integrating frameworks like Explainable AI (XAI) could bridge this gap, fostering trust and accountability in AI-aided storytelling processes.

To advance this field, future studies could explore the integration of multimodal sentiment analysis, combining text with visual and audio data to capture richer emotional cues, while also addressing ethical challenges such as algorithmic bias and data privacy. Additionally, investigating adaptive human-AI collaboration frameworks—where AI handles data processing and humans refine narratives—could optimise the balance between automation and creativity in data storytelling. Another promising avenue is examining the impact of cultural and linguistic diversity on sentiment analysis accuracy, alongside developing more inclusive AI tools to mitigate biases and improve cross-cultural applicability. Longitudinal studies assessing the real-world effectiveness of AI-enhanced dashboards in organisational decision-making would further validate their practical utility.

ACKNOWLEDGEMENT

This paper is financed by National Funds provided by FCT- Foundation for Science and Technology through project UID/04020: Research Centre for Tourism, Sustainability and Well-being.

REFERENCES

Abbass, H. (2021). Editorial: What is Artificial Intelligence? *IEEE Transactions on Artificial Intelligence*, 2(2), 94–95.

Anderson, E., & Zettelmeyer, F. (Eds.). (2021). *Leading with AI and Analytics: Build your data IQ to drive business value*. McGraw Hill.

Chen, W., Xu, Z., Zheng, X., Yu, Q., & Luo, Y. (2020). Research on sentiment classification of online travel review text. *Applied Sciences (Basel, Switzerland)*, 10(15), 52–75.

Costa, B. M. (2018). Mas afinal, para que serve a Análise de Sentimento? *YouTube*. <https://www.youtube.com/watch?v=JMQaFBYNLno>

Daradkeh, M., & Atalla, S. (2023). From Data to Narrative: Automating and Engineering the Art of Data Storytelling. In *International Conference on Information Technology* (Vol. 23, pp. 550-555).

Dykes, B. (2020). *Effective Data Storytelling: How to Drive Change with Data, Narrative and Visuals*. John Wiley & Sons.

El-Masri, M., Altrabsheh, N., & Mansour, H. (2017). *Successes and challenges of Arabic sentiment analysis research: a literature review*. Qatar University.

Ergen, M. (2019). What is Artificial Intelligence? Technical Considerations and Future Perception. *The Anatolian Journal of Cardiology*, 22, 5–7.

Feigenbaum, A., & Alamalhodaie, A. (2020). *The Data Storytelling Workbook*. Routledge.

Harari, Y. N. (2011). *Sapiens: História Breve da Humanidade - De Animais a Deuses*. (7ª ed.). Elsinore.

Hutto, C. J., & Gilbert, E. (2014). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. In *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media*. Association for the Advancement of Artificial Intelligence, 216-225.

Jyoti, R. (2021). Unlock The True Power of Data Analytics with Artificial Intelligence. In Liebowitz, J. (Ed.), *Data Analytics and AI* (pp. 21–30). CRC Press.

Knaflic, C. N. (2015). *Storytelling with data: A data visualization guide for business professionals*. John Wiley & Sons.

Li, H., Wang, Y., & Qu, H. (2024). *Where Are We So Far? – Understanding Data Storytelling Tools from the Perspective of Human-AI Collaboration*. The Hong Kong University of Science and Technology.

Liu, Y. (2021). Natural Language Processing in Data Analytics. In Liebowitz, J. (Ed.), *Data Analytics and AI* (pp. 117–131). CRC Press.

Lo Duca, A. (2023). Towards a Framework for AI-Assisted Data Storytelling. In *Proceedings of the 19th International Conference on Web Information Systems and Technologies* (pp. 512-519).

Morais, D. M., Oliveira, V., Junger, A. P., & Facó, J. F. (2020). O Conceito de Inteligência Artificial Usado no Mercado de Softwares, na Educação Tecnológica e na Literatura Científica. *Revista Educação Profissional e Tecnológica*, 4(2), 98–109.

Nelson, J. H. (2021). Data Science Life Cycle. In Liebowitz, J. (Ed.), *Data Analytics and AI* (pp. 1–7). CRC Press.

Poeta, E., Ciravegna, G., Pastor, E., Cerquitelli, T., & Baralis, E. (2023). Concept-based Explainable Artificial Intelligence: A Survey. *Politecnico di Torino*.

Quan, H., Li, S., Zeng, C., Wei, H., & Hu, J. (2023). Big Data and AI-Driven Product Design: A Survey. *Applied Sciences (Basel, Switzerland)*.

KEY TERMS AND DEFINITIONS

Artificial Intelligence (AI): A scientific field focused on creating machines or systems capable of performing tasks that typically require human intelligence, such as learning, reasoning, and decision-making. AI is applied in areas like data analysis, automation, and problem-solving, transforming industries through advanced algorithms and computational power.

AI-Powered Analytics: The use of AI algorithms to analyse large datasets in real time, extract actionable insights, and automate decision-making processes. It combines machine learning and natural language processing to enhance data interpretation, enabling businesses to uncover trends and optimise strategies efficiently.

Data Storytelling: The process of transforming raw data into compelling narratives by combining analysis, visuals, and storytelling techniques. It simplifies complex information, engages audiences, and supports data-driven decision-making in fields like business, journalism, and public policy.

Data-Driven Decisions: Decisions based on evidence derived from data analysis rather than intuition. By leveraging analytics tools, organisations identify patterns, predict outcomes, and implement strategies with greater accuracy, improving efficiency and competitive advantage. **Marketing:** Strategies and techniques used to promote products, services, or brands by addressing consumer needs. Modern marketing relies on data and AI to personalise campaigns, optimise engagement, and measure effectiveness across digital and traditional channels.

Sentiment Analysis: A natural language processing technique that detects emotions, opinions, and attitudes in text (e.g., social media comments). It classifies sentiment as positive, negative, or neutral, helping businesses gauge public perception and improve customer experiences.

Visual Storytelling: An approach that uses visual elements (charts, infographics, videos) to communicate data-driven stories. It enhances comprehension and retention by blending aesthetics with narrative, making complex information accessible and engaging for diverse audiences.

