Automated News in Brazilian television: a case study on the AIDA system (Globo-Brazil)

Automatización de textos periodísticos en la televisión brasileña: estudio de caso del sistema AIDA (Globo-Brasil)

ISSN: 1696-019X / e-ISSN: 2386-3978

Renato Essenfelder. PhD in Communication Sciences from the University of São Paulo (ECA-USP), Master in Portuguese at PUC-SP and Bachelor of Journalism. Professor of journalism at ESPM-SP since 2011, also a researcher at the Master of Journalism at the same institution since 2016, where he coordinates the Research Group on Content Production. His research work focuses in the areas of journalistic narratives, storytelling and Artificial Intelligence. He is also a columnist for "Portal Estadão".
Escola Superior de Propaganda e Marketing (ESPM), Brasil
renato.essenfelder@espm.br.
ORCID: 0000-0002-0618-1602

João Canavilhas. PhD at the Universidad de Salamanca (Spain) with a thesis entitled "Webnoticia: propuesta de modelo periodístico para la WWW". He is associate professor at the Universidade da Beira Interior (Covilhã – PORTUGAL), where actually is vice-rector and researcher at Labcom.IFP – Communication, Philosophy and Humanities. João Canavilhas is the author or co-author of 10 books, 36 chapter books and 45 papers in national and international scientific journals. His research work focuses on various aspects of "Communication and New Technologies", particularly in the fields of online journalism, social media and journalism for portable devices.
Universidade da Beira Interior, Covilhã, Portugal
jc@ubi.pt
ORCID: 0000-0002-2394-5264

Haline Costa Maia. Communication Specialist professional holding a Postgrad in Marketing, currently studying an MSc in Journalism at Universidade Fernando Pessoa (Portugal) and overall 13 years of experience in multinational companies such as Exxomobil and Accenture with planning, PR, movie screenwriting, copywriting, cinema, social media management, digital marketing and ad campaigns. Currently a Communication and Marketing Specialist at RPC - Rede Paranaense de Comunicação. Her preferred areas in academic research are NLP, artificial intelligence, communication, linguistics and technology.
Universidade Fernando Pessoa, Porto, Portugal
37617@ufp.edu.pt
ORCID: 0000-0003-0464-8616

How to cite this article:
Essenfelder, R.; Canavilhas, J.; Maia, H. C. and Pinto, R. J. (2019). Automated News in Brazilian television: a case study on the AIDA system (Globo-Brazil). *Doxa Comunicación*, 29, pp. 255-274.
https://doi.org/10.31921/doxacom.n29a13

Ricardo Jorge Pinto. Professor of Communication at the Fernando Pessoa University, journalist at Lusa news agency and commentator on National Politics at RTP (Public Broadcast Television Station). He has a PhD in Media Studies by the University of Sussex (UK). As a journalist he has worked at several television, radio and newspapers, as reporter, editor and editor-in-chief. His preferential area of journalistic work is political information. His preferential areas of academic research are Journalistic Techniques, Ideologies of Media, Political Journalism, and Digital Communication.
Universidade Fernando Pessoa, Porto, Portugal
ricpinto@ufp.edu.pt
ORCID: 0000-0001-5557-0690

ISSN: 1696-019X / e-ISSN: 2386-3978

Received: 27/07/2019 - Accepted: 12/11/2019

Abstract:

Technological advancements have created a media ecosystem in which traditional journalism sees its existence strongly threatened by the emergence of new players. Social networks have created a competitive environment that, whether due to its dispersion or its capillarity, has relegated the mainstream media to a secondary role in the media ecosystem. Ironically, the technologies that threaten traditional journalism are also those that can save it; provided they are used correctly. Journalism, weakened by the economic crisis and with increasingly smaller newsrooms, has artificial intelligence as an opportunity to recover a certain centrality in the media ecosystem. This paper studies AIDA, a project from the Brazilian television network Globo. This project looked to automation as a way to avoid errors and ambiguities in the news. The study of the AIDA case, complemented by interviews, presents the challenges to achieve the automatization of news regarding electoral polls.

Keywords:

Journalism, automation, Artificial intelligence, innovation, AIDA.

Recibido: 27/07/2019 - Aceptado: 12/11/2019

Resumen:

Los avances tecnológicos han creado un ecosistema mediático en el que el periodismo tradicional ve su existencia fuertemente amenazada por el surgimiento de nuevos actores. Las redes sociales han creado un entorno competitivo que, ya sea por su dispersión o por su capilaridad, relegaron los mainstream media a un rol secundario en el ecosistema mediático. Irónicamente, las tecnologías que amenazan el periodismo tradicional son asimismo aquellas que pueden salvarlo, siempre que se utilicen correctamente. El periodismo, debilitado por los efectos de la crisis económica y con redacciones cada vez más reducidas, tiene en la inteligencia artificial una oportunidad de recuperar cierta centralidad en el ecosistema de los medios. En este trabajo se estudia AIDA, un proyecto de la televisión brasileña Globo. Este proyecto buscó en la automatización una respuesta para evitar errores y ambigüedades en las noticias. El estudio del caso AIDA, complementado con entrevistas, presenta los desafíos para lograr la automatización en la lectura y presentación de datos de sondeos electorales.

Palabras clave:

Periodismo, automatización, Inteligencia Artificial, innovación, AIDA.

1. Introduction

The field of Artificial Intelligence (AI) attracts scientists from all areas. The new "golden age" of Artificial Intelligence (Havenstein, 2005), which began in the new century, is connected to the advancement of big data technologies that have provided alternatives to the slow and complex engineering of manually coded knowledge (Russel & Norvig, 2016), opening a wide range of opportunities in all sectors. As inventor and futurist Raymond Kurzweil (2005) said, "Today, many thousands of AI applications are deeply embedded in the infrastructure of every industry. Most of these applications were research projects ten to fifteen years ago" (p. 206). In spite of this popularity, defining AI is not a simple task (Brennen, Howard & Nielsen, 2018; Grant, Seruwagi, & Dodd, 2018; Obozintsev, 2018). The question "What is artificial intelligence?"

is one that has been asked for decades, with answers that vary according to the perspective adopted (Martínez-Plumed *et al.* 2018; McCarthy *et al.* 2006; Moor, 2006).

The definitions of AI can be divided into two large groups: those who have human beings as a reference (in other words, human imitative intelligences) and those who have abstract references such as "rationality" or "efficiency", without human intelligence being the central parameter (Russell & Norvig, 2016).

For the human paradigm group, AI can be the automation of "activities that we associate with human thinking, activities such as decision-making, problem solving, learning" (Bellman, 1978, p. 12) or, in the well-known formula by Kurzweil (1990), "The art of creating machines that perform functions that require intelligence when performed by people" (p.117).

The rationalist perspective (which does not have humans as a paradigm) has different definitions of AI, such as "the study of computations that make it possible to perceive, reason, and act" (Winston, 1992, p. 5), or as "intelligent behavior in artifacts" (Nilsson, 1998, p. 1). Also in this rational agent line, Russell & Norvig (2016) define AI as "the study of agents that exist in an environment and perceive and act" (p. 7). The rational agent approach, more common among engineers in this area, has helped drive the strong growth of AI in recent decades. In fact, it is more difficult to build machines that perfectly mimic a human being (as proposed by the English mathematician Alan Turing in his famous "Imitation Game"), than to build machines that solve complex problems; or as Russell & Norvig (2016) point out:

> AI researchers have devoted little effort to passing the Turing Test, believing that it is more important to study the underlying principles of intelligence than to duplicate an exemplar. The quest for "artificial flight" succeeded when the Wright brothers and others stopped imitating birds and started using wind tunnels and learning about aerodynamics. Aeronautical engineering texts do not define the goal of their field as making "machines that fly so exactly like pigeons that they can fool even other pigeons" (2016, p.3)

What we simply refer to as AI covers a large number of disciplines and can be observed in a wide range of applications. According to Russell & Norvig (2016), some of the main areas that explore AI are: a) robotic vehicles; b) voice recognition; c) autonomous planning and programming; d) games; e) spam control; f) logistic planning; g) robotics; and h) machine translation.

Likewise, "journalism is not apart from the process of global labor automation as a result of the improvement of artificial intelligence, robotics and new communication technologies, but it is argued that the tasks that require cognitive skills are more difficult to frame in standardized actions reproducible by a machine" (Túñez-López *et al.* 2018, pp. 756-757).

In journalism, artificial intelligence applications generally focus on the fields of Machine Learning and Natural Language Processing, which increasingly includes the automatic written texts to voice and vice versa, and robotics (Marconi & Siegman, 2013), but examples are still scarce. In an AI study specifically applied to investigative journalism, Stray (2019) pointed out that the examples in this regard are much rarer than what usually happens within the innovation and technology discourse adopted by media companies.

In this sense, this paper studies an application of artificial intelligence developed by the largest Brazilian television station, Rede Globo, to report on the results of electoral surveys. Through interviews with the creator of the AIDA system (Data

ISSN: 1696-019X / e-ISSN: 2386-3978

Interpreter Algorithm) and the analysis of the texts generated by the tool, we seek to record the motivation of the company and the results measured in this technological innovation movement in the journalistic field.

2. Artificial Intelligence, Automation and Algorithms

ISSN: 1696-019X / e-ISSN: 2386-3978

In order for this paper to achieve its objective, we must place artificial intelligence technologies within the wider universe of automation. Although all forms of AI involve some degree of automation (since the intelligent agent is able to react to environmental stimuli autonomously), not all automation requires AI. This distinction is necessary and relevant in the field because many AI engineers have recently complained about corporate marketing departments and their misappropriation of the term, betting on the fact that the mere mention of the term "artificial intelligence" would benefit the business.

According to a recent study by MMC Capital Fund in London, 40% of new European companies that claim to use artificial intelligence technologies in their businesses simply do not apply any digital technology of task execution that normally requires human intelligence (Olson, 2019). That same study goes on to say that there is an obvious reason as to why companies misrepresent AI: the projects that use this term in their business plans attract between 15% and 50% more funds than those that do not; investors do not like it if companies are unsure of the effective adoption of artificial intelligence (Everett, 2017).

This is because, in many cases, relatively simple automation tasks are confused with AI applications because the two technologies are both able to use algorithms to perform their tasks. In its most popular definition, algorithms are compared to recipes or to sets of instructions to follow to obtain a desired result (Bertocchi & Corrêa, 2012; Bruckner, 2018; Goffey, 2008). Algorithms may or may not be classified as "intelligent" depending on their mode of action; however, most algorithms are extremely simple (extremely straightforward), as Tutt (2017) noted. In such cases, the instruction set is quite limited and the result is easy to determine. Bruckner (2018) offers a good example of this: the algorithm that determines what the largest value is in a column of numbers in the Microsoft Excel program follows three elementary instructions: search the column, sort the numbers in descending order, and return the first result.

Even algorithms considered to be extremely complex (such as Google Page Rank, which organizes the results of any search) can have relatively simple mechanics that are more impressive because of the sheer large amount of data they process than for the architecture of this processing (Tutt, 2017) .

But the algorithms that pique the interest (and fears) of contemporary society the most are not those that follow fairly objective and predictable orders (referred to as "dumb algorithms") but a new type of algorithm called "learning", "artificial intelligence" or simply "intelligent". Intelligent algorithms, which have grown exponentially in recent years, are programmed to learn how to solve a task. Instead of creating a Page Rank to classify the Search results, they create a Page Rank to discover the best way to rank search results on the web. These learning algorithms are used in some the most impressive current AI developments, such as Tesla self-driving cars, the IBM Watson (which has beat humans in quiz shows) personal assistants such as Siri, Apple and Google Assistant, and facial recognition systems widely used in China as means of payment and even punishment for legal infractions, etc.

While the first algorithms are considered "automated", only the second ones (which are more autonomous and capable of reacting to changes in data entry and modifying the programming itself) can be considered truly "intelligent."

3. Automation and AI applications in journalism

Discussions about innovation in journalism have taken center stage in recent decades due to the rise of the Internet and the social transformations that have challenged the traditional business models of journalistic companies around the world. Given this new scenario, innovation in journalism has become fundamental (Küng, 2013) or even "the key to the viability of news media in the digital age" (Pavlik, 2013, p.190).

Journalism in the 21st century lives in a state of perpetual change (Nielsen, 2012) and relies on new digital technologies, especially those related to mobile devices (Küng, 2015). While investment in digital technologies has grown within news organizations in recent years, most corporations, seduced by expressions such as Artificial Intelligence, Virtual Reality, and Blockchain, have not presented a consistent innovation strategy, and thus have distanced themselves from their audiences (Posetti, 2018). In this era of perpetual change, the "Shiny Things Syndrome" led media companies to invest in the latest technologies and generated, in addition to financial losses, a kind of fatigue due to innovation (innovation fatigue) (Posetti, 2018).

While analyzing investments in technological innovation in journalism is important from a critical point of view, it is also undeniable that new technologies, when properly applied, have opened new possibilities and perspectives in journalism, either by making them cheaper, allowing previous knowledge that was previously unthinkable, or freeing professionals from exhaustive tasks, such as analyzing tens of thousands of documents.

In many cases, the most successful innovative technologies incorporated by media companies are invisible to the general public. One of the first large companies to add automation methods to their journalistic routine was the Associated Press, which automated the production of corporate financial news in 2015 (Graefe, 2016). In the same year, the World Editors Forum classified automated journalism as the main technological trend in newsrooms.

Since then, the application of algorithms, whether intelligent or not, has been extended to other communication companies to generate or support the production of news. The advantages are obvious: once developed, and after the investment stage, algorithms can produce faster, cheaper and potentially more accurate content than human journalists can (Graefe, 2016).

The adoption of algorithms in newsrooms led to the creation of the term "Automated Journalism", defined as "the process of using software or algorithms to automatically generate news stories without human intervention - after the initial programming of the algorithm, of course" (Graefe, 2016, p. 14). These algorithms applied to the journalistic routine allow for the automation of several stages of work such as data collection, subsequent analysis, writing, and the publication of texts.

Without restraint, artificial intelligence technologies have limitations in the production of journalistic content that can rarely be overcome, even over the medium term, by the human relationship component they presuppose. As Diakopoulos

ISSN: 1696-019X / e-ISSN: 2386-3978

ISSN: 1696-019X / e-ISSN: 2386-3978

(2019) explains, listening and reacting to testimonials in real time or dealing with sources of information are basic and unknown journalism tasks for AI. A recent report by the McKinsey Global Institute (2017) highlights the ability of robots and computers to acquire increasingly complex and sophisticated cognitive skills, but recognizes that AI in journalism can only cover 15% of reporters' tasks and 9% of editors' tasks.

Journalistic work has always been able to adapt to new technological ages and assimilate new features into its procedures. The telegraph, digital photography or computers are some examples of technologies that have helped journalism, although in some cases they have also reduced the importance of some traditional functions. The same goes for artificial intelligence technologies, which are changing some tasks within journalism. The Associated Press system mentioned above introduced robotic techniques able to identify thousands of photographs every day, and has given publishers more time to think about what they want to publish by removing the burden of sorting, which is quite time-consuming in itself. For Diakopoulos (2019), this example points to the possibility of an AI journalism where human tasks are hybridized, combined with algorithms that adapt to the capabilities of Artificial Intelligence and accommodate their limitations. As Salazar (2018) points out, "journalism is an area that will be significantly affected by the evolution of AI-related technologies, which unquestionably will change the way of dealing with the profession. The key is to move forward together. Robots and humans (p.311).

A precondition for the proper functioning of this type of automation is the existence of clear and well-structured data to power these systems. This is one of the main aspects that led "Automated Journalism" to be initially applied within the fields of sports and economics, where there are a lot of data for computers to analyze and the text is easier, since it is strictly referred to the communication of economical results or matches.

Another advantage for automating part of the news is the supposed impartiality attributed to the algorithm, an impartiality that only occurs if it was programmed neutrally. This is, by the way, the main peculiarity by the adoption of automation in more complex coverage, such as politics, as we will see in the case study of the AIDA system, developed within the Globo Group for the 2018 Brazilian presidential elections.

4. The 2018 presidential elections in Brazil

The 2018 Brazilian elections, held on October 7 (round 1) and October 28 (round 2), were one of the largest polarizing events in Brazilian society (Abranches, 2019; Almeida, 2019; Souza & Viscarra, 2019). For the first time in history, the favorite candidate for the post, former president Lula, affiliated with PT (Partido dos Trabalhadores), was arrested. Another candidate, Jair Bolsonaro, was stabbed during a campaign event. Bolsonaro would go on to be elected with 55% of the votes.

While polarization and the dissemination of false news was increasing on social networks such as Facebook, Twitter, and WhatsApp (Mergulhão *et al.*, 2018), the traditional press was losing credibility. In this scenario, Rede Globo, the largest television station in Brazil and one of the 20 largest media groups in the world, according to the consultant Zenith (2017), was criticized when the host of its nightly news program, "Jornal Nacional" (the largest audience in the

country), read a technically incorrect text about the results of the presidential election poll. Reading voting results from the Datafolha Institute, journalist William Bonner announced on September 14, 2018:

> Jair Bolsonaro of PSL had 22% in August, oscillated within the margin of error to 24%, now up to 26%. With a margin of error of 24 to 28. Ciro Gomes of PDT had 10%, then 13, now he remained at 13, by the margin of error, 11 to 15. Fernando Haddad from PT, had 4, then grew to 9, and now oscillates around 13. With the margin of error from 11 to 15 (Jornal Nacional, 2018, emphasis added).[1]

Criticisms focused on the misuse of the verb "oscillate." Technically, "oscillating" candidates' percentage of the votes varies within the margin of error of the survey. When the result exceeds the margin of error, it is correct to use expressions such as "up", "down" and other synonyms. Taking into account that the subject of the news had a margin of error of two percentage points, the variation from 9 points to 13 points for the *Partido dos Trabalhadores* candidate, Fernando Haddad, was in fact a growth, an increase in votes, and not a technical swing.

The reaction to this statement was immediate on social networks, to the point that exactly nine minutes after announcing the results of this survey, journalist William Bonner made a public correction, in which he has stated, verbatim:

> Let me make a correction. Not long ago, when we announced the Datafolha survey, we said that the PT candidate, Fernando Haddad, oscillated from 9% to 13%. According to Datafolha, the growth has occurred outside the margin of error, so the correct phrase is "the candidate Fernando Haddad has grown from 9% to 13%." We apologize for this mistake (National Journal, 2018).

This may seem like a technical detail, but considering the sensitivity of the issue and the fact that it was a mistake which misrepresented a candidate for the top position, it was enough to trigger criticism over Rede Globo. The case has led the Globo Group Research and Development team (MediaTechLab) to expedite an ongoing project: a system for generating texts to read on television news from electoral survey data. In 2015, Globo's technology research and development areas, which had been separate up until that time, merged into the "Pesquisa & Desenvolvimento centre" (Research and Development Centre). In 2019, this team was consolidated into one work group and was renamed MediaTechLab, its objective is to extend the limits of what can be done with media by combining communication, business and technology.

5. Methodology

The so-called fake news was at the center of the debate during the presidential elections in Brazil. Wardle (2017) proposes a scale for unreliable information that can range from "misinformation" (inadvertent dissemination of false information) to "disinformation" (intentional dissemination of false information). The objective of this investigation is not to analyze whether the Jornal Nacional case was random or planned, but rather to study the technology that Globo developed to respond to criticisms surrounding its biased transmission of information in the Brazilian elections

1 Original: "Jair Bolsonaro, do PSL, tinha 22% em agosto, oscilou dentro da margem de erro para 24%, agora foi para 26. Com a margem de erro, tem de 24 a 28. Ciro Gomes, do PDT tinha 10%, depois 13, agora manteve os 13, pela margem de erro, 11 a 15. Fernando Haddad do PT, tinha 4, depois cresceu para 9, agora oscilou para 13. Com a margem de erro tem de 11 a 15" (Jornal Nacional, 2018, grifos dos autores).

ISSN: 1696-019X / e-ISSN: 2386-3978

ISSN: 1696-019X / e-ISSN: 2386-3978

of 2018. By creating this technology, Globo tried to demonstrate that this was simply an example of "misinformation", corrected by the introduction of automation so that the information transmitted could become impartial.

For this study we used a mixed qualitative methodological approach composed of two parts: a case study complemented with interviews. This is an exploratory study as it aims to familiarize a problem in order to "make it more explicit or make assumptions. This main objective of this study is to improve ideas or discover intuitions"(Gil, 2002, p. 41).[2] Bervian, Cervo and Silva (2009, p. 63) reinforce that exploratory research aims to "become familiar with the phenomenon or obtain a new perception of it and discover new ideas."[3] The approach seemed appropriate to address an innovation tool which is still in the testing phase.

Interviews and case studies were used to support the inquiry of the subject. Haguette (2001, p. 86) defines the interview as "a process of social interaction between two people in which one of them, the interviewer, aims to obtain information from the other, the interviewee"[4]. For Yin (2010, p. 135), interviews are an essential source of evidence for case studies. He believes "well-informed interviewees can provide important insights. They can also provide shortcuts to the previous history of these situations, helping to identify other relevant sources of evidence"[5].

Interviews with members of the Rede Globo Research & Development team were conducted in three stages, starting with the most general (macro) to the most specific (micro). The general interviews were aimed at understanding the perception of innovation in journalism at TV Globo and about the investments made in it, as well as verifying operating Artificial Intelligence systems at the core of the Brazilian company journalism.

For this purpose, the first interview, conducted in February 2019, was based on a semi-structured questionnaire with questions organized into "innovation conception" groups; "innovation structure" (organization chart and team); and "innovation projects", as presented in Table 1.

2 Original: "Torná-lo mais explícito ou constituir hipóteses. Pode-se dizer que tais pesquisas têm como objetivo principal o aprimoramento de ideias ou a descoberta de intuições" (Gil, 2002, p. 41).

3 Original: "Torná-lo mais explícito ou constituir hipóteses. Pode-se dizer que tais pesquisas têm como objetivo principal o aprimoramento de ideias ou a descoberta de intuições" (Gil, 2002, p. 41).

4 Original: "Processo de interação social entre duas pessoas na qual uma delas, o entrevistador, tem por objetivo a obtenção de informações por parte do outro, o entrevistado" Haguette (2001, p. 86).

5 Original: "Entrevistados bem-informados podem proporcionar insights importantes. Eles também podem fornecer atalhos para a história prévia dessas situações, ajudando a identificar outras fontes relevantes de evidência" (Yin, 2010, p. 135).

Table 1: Example of initial questionnaire

Cluster	Basic question
Conception	What does innovation mean for TV Globo? How does it apply to the specific context of journalism?
Structure	What investments have been made in human resources, facilities, products or processes to generate innovation in TV Globo's journalism?
Projects	What past, present and future projects have been developed by the TV Globo innovation team? Do any of these projects incorporate Artificial Intelligence?

Source: author's research.

The first interview for the study was conducted with Daniel Monteiro, manager of the "Research & Development" team at TV Globo, and lasted for approximately 30 minutes. In the following days, the information was aggregated with two emails containing additional questions about the same clusters. At this point, authors mapped the work of the station in relation to Artificial Intelligence systems, which had not been previously investigated.

In March 2019, a new round of interviews was conducted, this time with R&D executive supervisor Cassius Estrada, and was complemented by questions sent via email. The development of the "Numerologists" and "AIDA" systems were added to this second stage, both of which are based on Artificial Intelligence and both developed to support the work of journalists in the researched company.

The third stage of the work contained five interviews with Giulio Bottari, an engineer responsible for the initial development of the AIDA system. The interviews were conducted between April and August of 2019, complemented by five other emails addressing specific technical questions about the AIDA system. A questionnaire was not applied at this stage. The open questions were designed to better understand how the system was developed and how it operates

After the interviews were completed, we moved on to the case study. According to Gil (2002, p. 54), a case study "consists in the absolute and exhaustive study of one or a few objects, in such a way as to allow their extensive and detailed knowledge"[6]. The case study is a strategy that seeks to respond to the "how" and the "why", applied to cases where "the researcher has little control over events and when focusing on contemporary phenomena within the context of life real" (Yin, 1994, p. 10)[7], as is the case with this investigation.

The interviews were complemented with access to the operational screens of AIDA (the interfaces for system users) and the visualization of recorded episodes in which the system was put into practice. On three occasions during October 2018, the station activated AIDA, generating texts that had been read on the Globonews cable channel. We transcribed and analyzed the text generated on one of these occasions to observe the fluidity and correction of the automated text.

6 Original: "Consiste no estudo profundo e exaustivo de um ou poucos objetos, de maneira que permita seu amplo e detalhado conhecimento" (Gil, 2002, p. 54).

7 Original: "O investigador tem pouco controlo sobre os acontecimentos, e quando o foco está nos fenómenos contemporâneos dentro do contexto da vida real" (Yin, 1994, p. 10).

ISSN: 1696-019X / e-ISSN: 2386-3978

The object of study, the AIDA system (Data Interpreter Algorithm), was designed in a short period of time (40 days) to respond to a problem that required a rapid answer. AIDA was developed in the context of the presidential controversy mentioned in the previous chapter and aimed to calculate the evolution of presidential candidates, detect the connections between them, and generate a text considered to be technically accurate and impartial (Bottari, 2019).

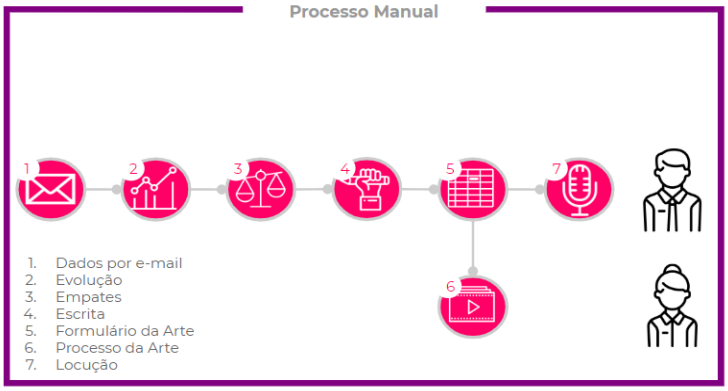
To understand how the system works, you must first understand the method commonly used in the TV Poll election routine. Bottari explains:

> Everything starts with a release from the research institute that arrives by email to the journalist. The release is large and contains a lot of data and questions that are asked of voters. The journalist (...) prints these sheets, which are nothing more than numbers and perhaps a graphic that accompanies them. If there are previous searches for other dates (for example, March, April and May), these previous results are also included in the release. The journalist takes the printed paper and manually calculates the margin of error for each date. Doing this will allow you to find out if the candidate: (a) maintained the index (that is, the same value as before); (b) fluctuated within the margin of error; (c) grew; or (d) fell in the search. Then you have to calculate (only for the most recent date) the connections considering the margin of error. This step is more difficult than it seems. The problem is finding connections and describing them in a condensed format. For example, if A is linked to B, and A is linked to D, and C is linked to B, and C to D, it would be ideal to say that "A and C are linked to B and D". In addition, there are different types of connections: (a) connections with the exact value, (b) technical connections within the margin of error and (c) connections at the limit of the margin of error. Since all three are reported differently, the journalist must keep this in mind. All this is done manually. Once this information has been obtained, the text can be written and then reviewed. The next step is to complete an Excel spreadsheet, which was developed by the Art team. The spreadsheet identifies the candidates and the data that the Art team will read, then generates the graphic in a semi-automatic process. The journalism team will then declare the text ready to be read. The audio is then recorded and sent to editing. The graphic is also produced and forwarded to editing, where it will be mixed with the story audio. (Bottari, 2019).

There are seven steps of manual work that must be done between receiving the data and delivering the text that is to be read (Figure 1). Not only is this complex, but the process is carried out under intense time pressure so that the data can reach the newsroom shortly before the news is scheduled to begin.

ISSN: 1696-019X / e-ISSN: 2386-3978

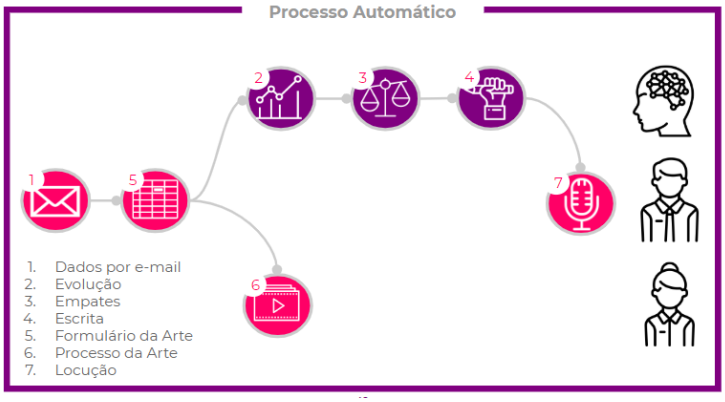
Figure 1: Manual text creation process for off reading



Source: Bottari, 2019.

AIDA was developed with the intention of discovering a way to save time, avoid errors, and minimize bias in reading the results. In the new system (Figure 2), the journalist receives the data by email (as before) and uploads them directly to a form. Because the completed form is used by another part of the workflow (graphics generation), the team did not develop a tool to extract data directly from the file sent by email. The system automatically calculates the evolution and the connections, and generates the text for reading. The journalist receives this text and forwards it to the voice over. With AIDA, the manual work phases are reduced from seven to four and, according to Bottari, the system saves up to 370 hours of labor for journalists in the new workflow (Bottari, 2019).

Figure 2: AIDA automated process

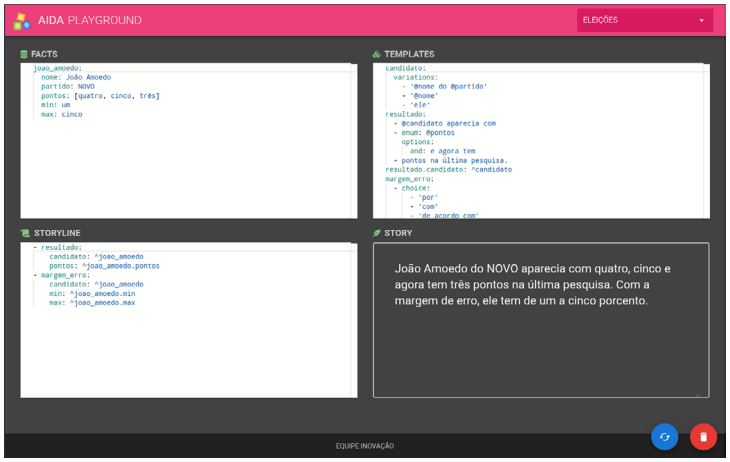


Source: Bottari, 2019.

ISSN: 1696-019X / e-ISSN: 2386-3978

The algorithm was developed in Python because the ease of programming in this language is suitable for creating prototypes quickly and intuitively, in addition to being widely used for data analysis and data science in general (Bottari, 2019). Figure 3 shows the interface visible to the journalist, with the text in the "story" window.

Figure 3: Interface showing the generation of text with electoral data



Source: Bottari, 2019.

The system was used three times during the elections, all for the Globonews cable channel, as it was considered unwise to test it on its main news channel (National Journal), despite the fact that it was the mistake made in that news program that led to the project being developed. According to Bottari (2009), journalists collaborated in designing the system in order to help "understand the rules they use, the correct words, etc.; that is, to clarify aspects about the values of news and journalistic impartiality for technicians, which shows the need for training teams to be diversified in their professional training.

6. Results and Discussion

As previously mentioned, the system was used three times during the elections. In this paper we analyze only one of the units because they are all the same in the problem generated by the development of the AIDA tool. In this case, it is the news issued in "Jornal Globonews" of October 3, 2018, where the journalist made the following narrative of AIDA:

> In the first survey, on August 24, Ibanês of MDB had 2%. It rose to 7, to 9, to 20, and now to 34%. The margin of error has it between 31 and 37%. Eliana Pedrosa of PROS appeared with 14% of the votes. It rose to 23. Then it oscillated around 22, 21, and now it is 17. The margin of error has it between 14 and 20%. Rodrigo Rollemberg of the PSB was 12%. He remained at 12 and then dropped to 11. He stayed at 11. And now he has 10%. It varies between 7 and 13. Alberto Fraga of DEM had 10%. 13.

ISSN: 1696-019X / e-ISSN: 2386-3978

> 14. 11. And now 9%. It ranges between 6 and 12. Technically it is linked with Rollemberg. Rogério Rosso of the PSD had 8%. 10. Maintained 10. It fluctuated to 11 and is now 8%. The margin of error has it between 5 and 11. Rosso is technically linked with Rollemberg and Fraga. General Paulo Chagas of the PRP had 3% in the first two surveys. Then 5, then 3, and now 4%. The margin of error has it between 1 and 7%. It is technically tied with Fraga and Rosso and on the edge of the margin with Rollemberg. White and null: 31%, 15, 13%, 11%, and now 8%. Those who did not know or did not vote: 12%, 8%, 8% again, and 4%. And now it is also 4%.(Globo News in Pauta, 2018).[8]

It is interesting to observe how the system clearly evaluates the cases of a tie between the candidates. The system perceives the difference between "they are tied" (that is, they registered the same absolute number of votes), "they are technically tied" (tie considering the margin of error, in this specific case, of three percentage points) and "they are tied at the margin of error" when there are six nominal points separating the candidates. The system also evaluates how many candidates are simultaneously in a tie, that is, it not only inc'icates the tie between X and Y, but also between X, Y and Z, respecting the different voting rates, which requires a more complex level of automation.

Once the structure of the text is analyzed, the texts generated by the computer and those generated by humans are practically indistinguishable in terms of structure and vocabulary (Table 2).

ISSN: 1696-019X / e-ISSN: 2386-3978

8 Original: "Na primeira pesquisa, em 24 de agosto, Ibanês, do MDB, tinha 2 %. Foi para 7, oscilou para 9, subiu para 20 e agora cresceu para 34%. Pela margem de erro está entre 31 e 37. Eliana Pedrosa, do PROS, aparecia com 14% das intenções de voto. Cresceu para 23. Depois, oscilou para 22, 21, e agora aparece com 17. Pela margem de erro tem entre 14 e 20%. Rodrigo Rollemberg, do PSB, tinha 12. Manteve os 12. Depois 11. Continuou com 11. E agora tem 10. Varia entre 7 e 13. Alberto Fraga do DEM, tinha 10%. 13. 14. 11. E agora aparece com 9%. Oscila entre 6 e 12. Ele está tecnicamente empatado com Rollemberg. Rogégio Rosso, do PSD tinha 8%. 10. Manteve os 10. Oscilou para 11 e agora está com 8%. Pela margem de erro fica entre 5 e 11. Rosso está tecnicamente empatado com Rollemberg e Fraga. General Paulo Chagas, do PRP, tinha 3 nas duas primeiras pesquisas. Depois 5. 3. E agora tem 4%. Pela margem de erro fica entre 1 e 7%. Ele está tecnicamente empatado com Fraga e Rosso e no limite da margem com Rollemberg. Brancos e nulos: 31. 15. 13. 11 e agora 8%. Não souberam ou não responderam: 12. 8. 8 de novo. 4. E agora também são 4 %." (Globo News em Pauta, 2018). Disponible em: http://g1.globo.com/globo-news/globo-news-em-pauta/videos/t/todos-os-videos/v/ibope-divulga-pesquisa-de-intencao-de-voto-para-o-governo-do-distrito-federal/7063121/

ISSN: 1696-019X / e-ISSN: 2386-3978

Table 2. Human text vs. computer generated text

	Human generated	Computer generated
Verbal text (original excerpt from the first three in the surveys)	Jair Bolsonaro, from PSL, had 22% in August, oscillating within the margin of error of 24%, and now with 26%. The margin of error is between 24 and 28%. Ciro Gomes, from PDT, had 10%, then 13, and now remains at 13. The margin of error is between 11 and 15%. Fernando Haddad, from PT, had 4%, then rose to 9%, and now oscillates around 13%. The margin of error is between 11 to 15%. Ciro Gomes and Fernando Haddad are still tied.	The first study on August 24, Ibanês from MDB had 2%. He rose to 7%, oscillated around 9, rose to 20, and ios now at 34%. The margin of error is between 31 and 37%. Eliana Pedrosa, from PROS, appeared with 14% of votes. She rose to 23, and then oscillated around 22, 21, and now has 17. The margin of error is between 14 and 20%. Rodrigo Rollemberg, from PSB, had 12%. He remained at 12. Then dropped to 11. Continued with 11. Now he has 10%. The margin of error is between 7 and 13%.
Date of issue	14/9/2018	3/10/2018
Newscast	Jornal Nacional (TV Globo)	Em Pauta (Globonews)
Verbs	Have, rock, go, stay, rise, be.	Have, go, swing, climb, rise, be, appear, stay, continue.
Substantives	The month of August The margin of error <candidate name> <candidate party>	**The research** The month of August **Votes** The margin of error <candidate name> <candidate party>
Adverbs, adjectives		
Sentence structure	Subject-predicate, in direct order.	Subject-predicate, in direct order.

Source: Authors

As noted, both texts have a fairly simple structure in terms of vocabulary and syntax. They follow the same basic structure and do not vary much. The same set of verbs is repeated, synonyms or complex structures are avoided, as well as adjectives and adverbs. The second text, generated by computer, has an additional element of difficulty: it narrates a historical set of five surveys per candidate, compared to three surveys narrated in the text generated by humans. Perhaps this is why there is a slightly larger number of different verbs (9 as compared to 6 in the human text). The verbs that do not appear in the first text are "upload," "appear," and "continue." All are simple synonyms for other verbs used in both texts, such as "rise," "have," and "stay," respectively.

The second text was generated by the computer language AidaLang, which was created to facilitate the creation of text templates. "With it, you can define how tie templates will look, the evolution, the description of the candidate, etc. without having to change the Python code. This is the first step towards later creating a tool that will allow a journalist to create templates", thus giving the human editor more power to modify the AIDA output (Bottari, 2019).

According to the developer, it is not easy to classify AIDA as an Artificial Intelligence system, which is in line with literature on the subject which has difficulty conceptualizing the expression clearly. "What we have is a language (AidaLang) that facilitates the generation of natural language text" (Bottari, 2019). According to Bottari, "we can say that AIDA is not an AI system, but it is not a system of pure and simple rules since we have a heuristic rule for ties". The system lies somewhere between an AI and a simpler automation.

The heuristic rule indicated by the developer is based on graph theory[9], an important concept in NLP language programming. For the AIDA project, a model was developed which represents candidates as nodes and ties as edges, concepts that describe a possible pathway from A to B. The objective is to find a way to "compress" the ties so that the speaker has to read the least number of names when mentioning ties.

> For example, it is better to say "A and C are linked to B" (3 names) than to say "A is linked to C, and B is linked to C" (4 names). In the model, we start with a set of ties, for example, {A} -> {B} and {B} -> {C}. What heuristics does is try to reduce the number of nodes. For this, it performs the following tests:
> {A} -> {B} + {B} -> {C} (cannot join);
> {A} -> {B} + {C} -> {B} => {A, C} -> {B} (can be joined) (Bottari, 2019).

For the future, Globo researchers are developing an application which will further automate the workflow. The most important aspect of this application includes automated voice-over in a process known as text-to-speech conversion, which is part of the AI technology group known as NLP (Natural-language Processing). Here, the automated voice (read by a robot) is accompanied by a 3D animation made with graphic computing (Fig. 4). Bottari clarifies:

> In this application, voice synthesis processes are performed through the Microsoft Text-to-Speech service. Once synthesized, the voice is mapped into visemas (oral animations that correspond to the phonemes of the words) in real time. In turn, visemas can compose a real-time representation of a character (AIDA) that presents a newscast. The simulation works like a video game and was synthesized with the Unreal Engine 4 platform, a general application that allows you to generate more realistic characters (Bottari, 2019).

The MediaTechLab team first presented the computer graphics with the voices on September 11, 2019, at the IBC event - a technology fair held in the city of Amsterdam. Voice automation allows for a choice of colloquial or formal language tones, as well as the ability to choose different scenarios, and until now has never been used in Rede Globo television channels.

9 Graph theory is a branch of mathematics that studies the relationships between objects in a given set. It uses structures called graphs, with nodes and edges. Two nodes are connected if there is a path between them. (Biggs et al., 1986).

ISSN: 1696-019X / e-ISSN: 2386-3978

Figure 4. 3D modeling prototype

ISSN: 1696-019X / e-ISSN: 2386-3978

7. Conclusions

This paper analyzes the response from the largest Brazilian television network (Globo) to an event that affected this television network's image in the Brazilian 2018 presidential elections. It should be noted that the partiality, or impartiality, of Globo in these elections is not the object of study in this paper. We are fundamentally interested in the need to avoid human errors which typically occur in current journalism - the lack of time for producing rigorous news. Rede Globo decided to develop a news automation system (AIDA) to eliminate this problem, which is the real object of study in this paper.

The heuristics used to develop AIDA belong to a class of algorithms that, although do not guarantee the best and most compact solutions possible, are fast, light, easier to use and quite precise. In this case, it automatically determines the ties considering the margins of error in the surveys. It is interesting to note that mathematics acts as a resource for textual cohesion to suppress redundant words and elaborate sentences that seek the same natural tone of human journalistic language.

Using the AIDA project as a parameter, the inclusion of journalists in forming Artificial Intelligence tools that development teams applied to journalism is an important recommendation, as programmers complain of difficulties in understanding the daily challenges of news professionals and especially of values associated with the field, such as clarity, balance and impartiality, enhanced by the enormous exposure the texts are subjected to. What engineer would imagine that simply substituting the verb "rise" with the verb "oscillate" in an electoral news survey could lead to so many negative repercussions?

In addition to the semantics, the advancement of the possibilities of AI applications lead to questions related to the agenda, newsworthiness criteria and journalistic ethics, which require journalists to collaborate towards better

codification. In this sense, the word "collaboration" seems to be more important than the word "substitution" when it comes to integrating AI projects with journalism newsrooms. At least for now, in the current stage of these technologies, collaboration is indispensable.

In terms of future research in the academic field, introducing Artificial Intelligence tools and technologies into newsrooms leads to a wide range of studies on topics such as: 1) the effectiveness of these solutions, that is, their ability to add value to the journalist's job or make it less strenuous; 2) the risks these technologies portray, from an employment point of view, with the reduction of jobs, a social point of view, and the quality of information presented; 3) the processes of development and implementation of these solutions, and the internal conflicts they generate; 4) the associated ethical aspects, including the possibility that the human imitative systems confuse the audience (or even the sources of information, which can be interviewed by robots in the near future); 5) the financial aspects of this movement, with the possibility of large technology corporations taking control of journalistic companies, as has already occurred with the Washington Post being bought by Amazon.

There is also much to discuss about the context of training young journalists. The following question arose during this investigation at various times: Should journalists, who have advised on the development of AIDA, study programming in order to provide for better collaboration?

The current stage of AIDA, which is moving towards developing graphic interface, still raises many research questions. What are the consequences of automated voiceover narration, that is, systems that can mimic the voice of human journalists who are widely known to the public? This innovation raises an old fear: can machines replace humans in activities of human intelligence and sensitivity such as storytelling?

In the case of AIDA, voice synthesis is applied more to the concept of imitating human performance, of making viewers feel more comfortable than with texts read by a robotic voice. However, this technology opens up opportunities in which a newscast can be customized to the public's taste, in the same way that social networks are. In this case, it will be up to the public to not only state what news they want to hear, but also whose voice they want to hear, what the tone of that voice will be, and who those synthesized characters will be on mobile devices or on Internet-connected television. According to the forecasts of Globo researchers, the advancement of AI will one day automatically generate personalized content, something that is impossible for a human news anchor to do.

This evolutionary perspective does not even take into account the advances still to be made in the field of AI, which is in a new "golden age", constantly testing new possibilities and applications. For now, however, there are notable limitations on automation systems such as AIDA when applied to complex fields like political journalism. In this regard, it is important to point out the fundamental role of journalists in identifying the semantic contexts of the proposed automation, such as the incorrect use of the term "oscillate" that supported the development of this algorithm. Previously knowing the meaning of this word would not have been enough to understand the error it caused: it is necessary to understand that, in the context of political journalism, the word represents a small variation within a margin of error.

ISSN: 1696-019X / e-ISSN: 2386-3978

The same difficulty extends to the syntactic questions, when the active participation of journalists could represent other possibilities of construction and textual analysis through the data, although in the case of AIDA, the text's small complex language becomes attractive to Grupo Globo due to its obvious impartiality. With this in mind, resistance to these systems by journalists tends to bring fewer benefits than advantages, given the many possibilities that these technologies could offer to newsrooms. The AIDA system, although it has its limitations, shows that, when working together, the contributions of journalists and technology experts can achieve interesting results.

ISSN: 1696-019X / e-ISSN: 2386-3978

8. Bibliographic references

Abranches, S. (2019). *Democracia em risco? 22 ensaios sobre o Brasil hoje*. S. Paulo: Companhia das Letras.

Almeida, R. D. (2019). Bolsonaro presidente: conservadorismo, evangelismo e a crise brasileira. *Novos Estudos Cebrap*, São Paulo, v. 38, n. 1, p. 185-213.

Bellman, R. E. (1978). *An Introduction to Artificial Intelligence: Can Computers Think?* Boyd & Fraser Publishing Company.

Bervian, P. A.; Cervo, A. L.; Silva, R. (2007). *Metodologia científica*. Pearson Prentice Hall.

Bottari, G. (2019). Personal interview.

Brandão, A. G., Bottari, G. D., Rodrigo, Corbelli, C., & Frensch, T. (2019, May 10). Personal interview.

Brennen, J. S., Howard, P. N., & Nielsen, R. K. (2018). An Industry-Led Debate: How UK Media Cover Artificial Intelligence.

Bruckner, M. A. (2018). The promise and perils of algorithmic lenders' use of big data. *Chi.-Kent L. Rev.*, *93*, 3.

Biggs, N., Lloyd, E. K., & Wilson, R. J. (1986). *Graph Theory, 1736-1936*. Oxford University Press.

Corrêa, E. S., & Bertocchi, D. (2012). A cena cibercultural do jornalismo contemporâneo: web semântica, algoritmos, aplicativos e curadoria. *Matrizes*, *5*(2), 123-144.

Diakopoulos, N. (2019) *Automating the News: How Algorithms Are Rewriting the Media*. Harvard University Press, Cambridge (Mass.).

Everett, C. (2017). Is Ai Hype Causing Damaging Confusion? *Computer Weekly*, 25–29. Accessed on June 3, 2019 http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=126465920&lang=pt-pt&site=ehost-live.

Gil, A. C. (2002). *Como elaborar projetos de pesquisa*. São Paulo: Atlas.

Globonews Em Pauta. (2018, October 03). *Ibope divulga pesquisa de intenção de voto para o governo do Distrito Federal*. Accessed on May 17, 2019. http://g1.globo.com/globo-news/globo-news-em-pauta/videos/t/todos-os-videos/v/ibope-divulga-pesquisa-de-intencao-de-voto-para-o-governo-do-distrito-federal/7063121/

Goffey, A. (2008). *Algorithm*. Software Studies - A lexicon. MIT Press.

Graefe, A. (2016). *Guide to Automated Journalism*. Tow Center for Digital Journalism Report. Accessed on May 21, 2019. https://academiccommons.columbia.edu/doi/10.7916/D8QZ2P7C/download.

Grant, A. W., Seruwagi, L. Z., & Dodd, M. S. (2011). *Artificial Intelligence Through the Eyes of the Public*.

Haguette, T. (2001). *Metodologias Qualitativas da Sociologia*. Petrópolis: Vozes

Havenstein, H. (2005). *Spring comes to AI winter*. Computer World.

Küng, L. (2013). Innovation, technology and organisational change. *Media innovations: A multidisciplinary study of change*, 9-12.

Kurzweil, R. (1990). *The Age of Intelligent Machines*. MIT Press.

Kurzweil, R. (2005). *The Singularity is Near*. Viking.

Marconi, F., & Siegman, A. (2017). *The future of augmented journalism: A guide for newsrooms in the age of smart machines*. New York: AP Insights.

Martínez-Plumed, F., Loe, B. S., & Flach, P. Ó hÉigeartaigh, S., Vold, K., & Hernández-Orallo, J. (2018). The facets of artificial intelligence: A framework to track the evolution of AI. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence* (pp. 5180-5187).

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). *A proposal for the dartmouth summer research project on artificial intelligence*, August 31, 1955. *AI magazine*, *27*(4), 12-12.

McKinsey Global Institute (2017). *A Future that Works: Automation, Employment and Productivity*. Accessed on June 20, 2019. https://www.mckinsey.com/~/media/McKinsey/Featured%20Insights/Digital%20Disruption/Harnessing%20automation%20for%20a%20future%20that%20works/MGI-A-future-that-works_Executive-summary.ashx

Mergulhão, D. R. D. S., Júnior, M., Oliveira, J. C., & Albuquerque, P. F. (2018). Post-truth, fake news e processo eleitoral. *Revista de Estudos Eleitorais:* vol. 2, n. 4.

Moor, J. (2006). The Dartmouth College artificial intelligence conference: The next fifty years. *Ai Magazine*, *27*(4), 87-87.

Nielsen, R.K. (2012) *Ten Years That Shook the Media World. Big Questions and Big Trends in International Media Developments*. Report (Univ Oxford, Oxford).

Nilsson, N. J. (1998). *Artificial Intelligence: A New Synthesis*. Morgan Kaufmann.

Obozintsev, L. (2018). *From Skynet to Siri: an exploration of the nature and effects of media coverage of artificial intelligence* (Doctoral dissertation, University of Delaware).

Olson, P. (2019). Nearly Half Of All "AI Startups" Are Cashing In On Hype. *Forbes.Com*, N.PAG. Accessed on June 3, 2019. http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=135056985&lang=pt-pt&site=ehost-live

Pavlik, J. V. (2013). Innovation and the Future of Journalism. *Digital journalism*, *1*(2), 181-193.

Posetti, J. (2018). Time to step away from the 'bright, shiny things?' Towards a sustainable model of journalism innovation in an era of perpetual change. *RISJ*

Research Report. Oxford: University of Oxford.

Russell, S.J.; & Norvig, P. (2016). *Artificial intelligence: a modern approach*. Pearson Education Limited.

ISSN: 1696-019X / e-ISSN: 2386-3978

Salazar, I. (2018). Los robots y la Inteligencia Artificial. Nuevos retos del periodismo. Doxa Comunicación, 27, pp. 295-315. (https://doi.org/10.31921/doxacom.n27a15)

Stray, J. (2019). Making Artificial Intelligence Work for Investigative Journalism. *Digital Journalism*, 1-22.

Túñez-López, José-Miguel; Toural-Bran, Carlos; Cacheiro-Requeijo, Santiago (2018). Uso de bots y algoritmos para automatizar la redacción de noticias: percepción y actitudes de los periodistas en España. *El Profesional de la información*, v. 27, n. 4, pp. 750-758 (https://doi.org/10.3145/epi.2018)ul.04

Tutt, A (2017) An FDA for algorithms. *SSRN Scholarly* Paper, Rochester, NY: Social Science Research Network. Accessed on June 3, 2019. http://papers.ssrn.com/abstract=2747994.

Viscarra, S. P., & Souza, B. M. (2019). Democracia e Polarização Eleitoral no Brasil: as opiniões dos eleitores de PT e do PSDB (2002-2014). *Missões: Revista de Ciências Humanas e Sociais*, 4(3).

Wardle, C. (2017). Fake News. It's Complicated. *Medium.com*. Accessed on June 18, 2019. https://medium.com/1st-draft/fake-newsits-complicated-d0f773766c79

Winston, P. H. (1992). *Artificial Intelligence* (Third edition). Addison-Wesley.

Yin, R. K. (1994). *Pesquisa Estudo de Caso - Desenho e Métodos*. Porto Alegre: Bookman.

Yin, R. K. (2010). *Estudo de caso: planejamento e métodos*. Porto Alegre: Bookman.

Zenith. Top 30 Global Media Owners 2017. (2017, June 27). Accessed on June 2, 2019. https://www.zenithusa.com/top-30-global-media-owners-2017/

ISSN: 1696-019X / e-ISSN: 2386-3978