Article

Artificial Intelligence in News Media: Current Perceptions and Future Outlook

Mathias-Felipe de-Lima-Santos [1,*] and Wilson Ceron [2]

1 School of Communication, University of Navarra, 31009 Pamplona, Navarra, Spain
2 Institute of Science and Technology, Federal University of São Paulo, São Paulo 04021-001, Brazil; wilson.seron@unifesp.br
* Correspondence: mdelimas@unav.es

Abstract: In recent years, news media has been greatly disrupted by the potential of technologically driven approaches in the creation, production, and distribution of news products and services. Artificial intelligence (AI) has emerged from the realm of science fiction and has become a very real tool that can aid society in addressing many issues, including the challenges faced by the news industry. The ubiquity of computing has become apparent and has demonstrated the different approaches that can be achieved using AI. We analyzed the news industry's AI adoption based on the seven subfields of AI: (i) machine learning; (ii) computer vision (CV); (iii) speech recognition; (iv) natural language processing (NLP); (v) planning, scheduling, and optimization; (vi) expert systems; and (vii) robotics. Our findings suggest that three subfields are being developed more in the news media: machine learning, computer vision, and planning, scheduling, and optimization. Other areas have not been fully deployed in the journalistic field. Most AI news projects rely on funds from tech companies such as Google. This limits AI's potential to a small number of players in the news industry. We made conclusions by providing examples of how these subfields are being developed in journalism and presented an agenda for future research.

Keywords: journalism; artificial intelligence; computer science; machine learning; computer vision; NLP

1. Introduction

We are now living in a world where technology and data guide a large array of decisions in our lives. This means that every industry must adapt and embrace these technologies to become sustainable in the future. This is no different for the news industry since "the future of journalism and its business models seem to be bound by the penetration of technological deployment" (de-Lima-Santos and Mesquita 2021a, pp. 1416–17).

In recent years, news media has been greatly disrupted by the potential of technologically driven approaches in the creation, production, and distribution of news products and services (Hernandez Serrano et al. 2015; Örnebring 2010). This can be seen in novel news products and practices such as data journalism (Coddington 2015; Hermida and Young 2019), immersive and drone journalism (Harvard 2020; Kang et al. 2019), analytics (Nelson and Tandoc 2018; Ferrer-Conill and Tandoc 2018), and automation (Linden 2017b; Lewis et al. 2019).

Over the last several years, scholars have expanded their knowledge and experiences by confirming that artificial intelligence (AI) capabilities are evolving every year, which have made it cost less and offer more affordable computing power. The concept of AI refers more narrowly "to a branch of computer science focused on simulating human intelligence" (Broussard et al. 2019, p. 673).

AI in news media can certainly make journalism easier for overburdened resources without replacing journalists' unique skills. Furthermore, AI can enhance new forms of participation and leverage new products that could increase news media consumption

(Diakopoulos 2020; Jamil 2020). However, several challenges tend to delay technological innovations in newsrooms, such as the resistance to change, the institutional landscape, historical competition, insufficient funding, a lack of skill, and complementary ambitions (Boczkowski 2004; de-Lima-Santos and Mesquita 2021b; Krumsvik et al. 2019; Paulussen 2016).

Despite the practical deployment of AI in news media, the academic literature on this subject is still at a nascent stage. In this context, this study aims to expand the existing literature by examining case studies collected by JournalismAI, a project from the London School of Economics and Political Science (LSE). The list includes 102 cases that were collected by the institution and examples that were submitted by other researchers and practitioners in the AI field. Although this list is not exhaustive, it highlights the level of the development of AI in news and related areas.

To make a theoretical contribution to this field, this study draws upon literature on artificial intelligence to sketch an outline of the field and understand where the journalism industry is positioning itself. Thus, this article poses the following questions:

RQ1. How is news media positioning itself in the subfields of artificial intelligence?
RQ2. To what extent is AI being deployed in the news industry?
RQ3. What are the future avenues for AI in news media?

The contribution of this paper is twofold. First, it describes the types of AI that are being developed in the news media industry by facilitating research on the topic and helping practitioners to understand the possibilities, implications, and responsibilities that come along with AI adoption and use. Second, it provides a comprehensive analysis of the related subfields to encourage scholars and professionals to consider further steps for utilizing AI in the news industry.

2. Theoretical Framework

2.1. Artificial Intelligence in Its Current Manifestation

Although the term "artificial intelligence" has gained popularity in recent years, AI is not new. It dates back to 1955 when Stanford University's Professor John McCarthy used the term to describe the science and engineering of making intelligent machines. (McCarthy 1998). The pervasive nature of information and communications technology (ICT) and the datafication of society have expanded their applicability in a variety of fields, such as journalism (Gelgel 2020).

As technology evolves, new opportunities become available to consumers and businesses and help to address some of the challenges of our time (Paulussen 2016). As a result, the field of artificial intelligence has seen significant progress in recent years and has been led by numerous technological developments that have made it more affordable. Although Hollywood and science fiction movies often depict AI as sentient machines such as robots that can mimic human reasoning and behavior (Broussard 2018), the field is concerned with understanding and building intelligent entities that can compute "how to act effectively and safely in a wide variety of novel situations" (Russell and Norvig 2021, p. 19). Therefore, intelligence involves performing human tasks such as recognizing images or performing repetitive tasks (Broussard et al. 2019). In this sense, some authors understand intelligence in AI as rationality, which can be loosely understood as making correct decisions (Russell and Norvig 2021). However, scholars and experts have developed different definitions over the years. In simple terms, AI can be defined as the process of "creating computing machines and systems that perform operations analogous to human learning and decision-making" (Castro and New 2016, p. 2). Therefore, AI represents a step-by-step procedure for solving problems.

The academic scholarship of AI has been part of computer science for decades, but the automation of cognitive tasks became possible only after the developments relating to data, sensors, and advances in technology (Chan-Olmsted 2019). The ubiquity of computing has become apparent and has demonstrated the different approaches that can be achieved using AI. Based on opinions about the most promising methods and theories, experts have effectively deployed AI in some fields and have discovered that utilizing it in other fields is still unattainable (Aronson 2018; Castro and New 2016; Ortiz Freuler and Iglesias 2018).

Despite AI being a disputed concept, seven subfields emanate from this major field and share significant connections and commonalities: (i) machine learning; (ii) computer vision (CV); (iii) speech recognition; (iv) natural language processing (NLP); (v) planning, scheduling, and optimization; (vi) expert systems; and (vii) robotics.

The most popular of these subfields is machine learning, which is a "subfield of AI that studies the ability to improve performance based on previous experience" (Russell and Norvig 2021, p. 19). In other words, machine learning is a branch of AI that is dedicated to designing algorithms that build models from data without pre-existing solutions to a problem (Castro and New 2016). This has reduced operational costs and the cost of manpower substantially in the news industry.

The development of machine learning highly relates to deep learning and predictive analytics. In deep learning, statistical techniques are used to solve problems with little human intervention. To solve problems, models rely on large and complex datasets to replicate the human brain's learning capabilities (Chan-Olmsted 2019; Hassaballah and Awad 2020). The designs of some of these models were inspired by the structure and function of neural networks that were used to enable a computer to learn to recognize abstract patterns to simulate large, multilayered webs of virtual neurons (Castro and New 2016). Predictive analytics is a branch of machine learning that is dedicated to making predictions about future outcomes using historical data (Russell and Norvig 2021).

Another subfield of AI is computer vision. Through the use of mathematical algorithms, CV allows computers to derive meaningful information from digital images (Szeliski 2011). This does not mean that a CV model can actually see the content of an image like a human can, but it is capable of detecting or deducing an object (Marr 2010). CV algorithms can be divided into two strands: image recognition and machine vision. Image recognition involves a set of methods for detecting and analyzing images that can be used for the automation of specific tasks. Machine vision encompasses a computer's capabilities to perceive an environment (Szeliski 2011; Marr 2010).

Speech recognition focuses on automatically and accurately transcribing human speech and converting voice data into text data; it is commonly found in applications that follow voice commands or answer spoken questions (Deloitte 2014). Natural language processing goes beyond that and refers to the automatic computational processing of human language (Castro and New 2016). In other words, NLP is a computer program's ability to manipulate text and spoken words like humans do by understanding and responding to text or voice data, extracting meaning from sentences, or generating readable texts (Deloitte 2014). This broader subfield includes other areas such as translation, classification, and clustering, as well as information extraction. Other terms that relate to NLP are natural language generation (NLG) and natural language understanding (NLU), as illustrated in Figure 1. While the former is responsible for converting structured data into meaningful sentences in the form of natural language, the latter represents the process that turns unstructured data into understandable structured data (Future Today Institute 2018; Locker et al. 2019).
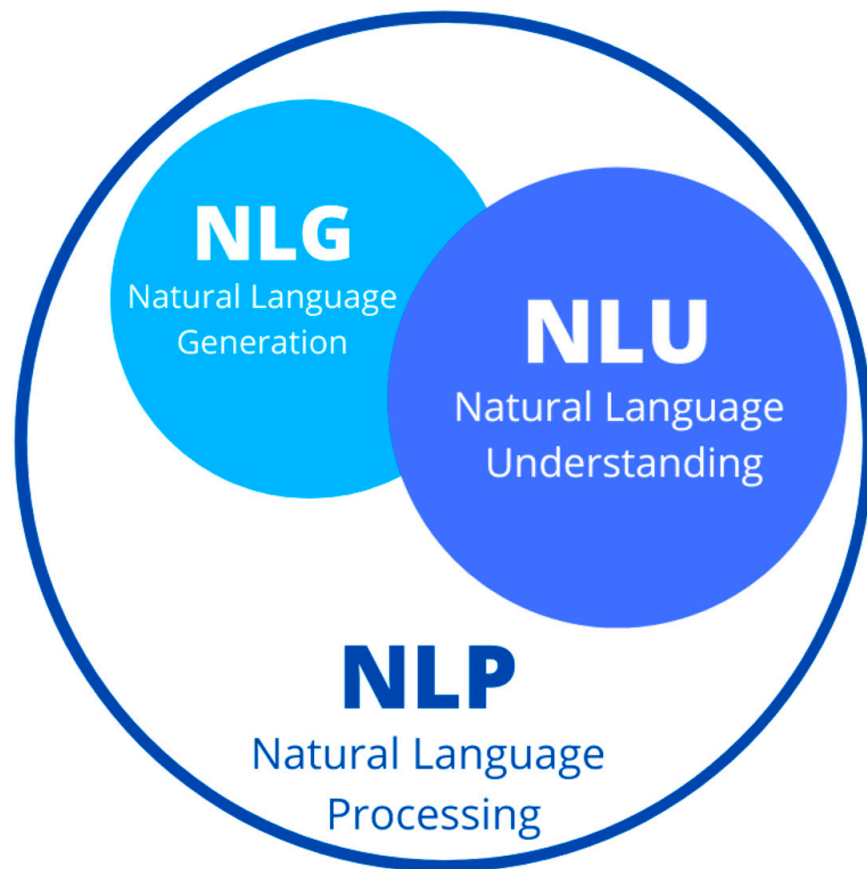
Figure 1. NLP models can be divided into two subsections: NLG and NLU (Source: Authors).

Other mature cognitive technologies include systems that use AI to determine steps to take (planning) and understand when to carry out a certain step (scheduling) to achieve a goal. Additionally, such a system can find the most optimal way to make all the necessary adjustments in the most efficient manner (optimization). This step is important in some cases due to trade-offs concerning limited resources and complex decisions required of the model. This subfield is known as planning, scheduling, and optimization (Deloitte 2014; Russell and Norvig 2021).

Rules-based systems simulate the behavior and judgment of humans who have expert knowledge and experience in a particular field to automate the process of making inferences about information. Known as expert systems, these models use databases of knowledge and rules to solve complex problems (Russell and Norvig 2021). Robotics is a subfield of AI that integrates different cognitive technologies to enable computers and systems to perform different tasks in conjunction with people in unpredictable environments. Some examples of robotics are robotic vacuums and unmanned aerial vehicles (Russell and Norvig 2021).

Although these subfields describe different applications of AI, they are interwoven and often mutually reinforcing. For this reason, some authors view AI according to its five functions: monitoring, discovering, predicting, interpreting, and interacting with physical environments, humans, or machines (Castro and New 2016).

However, artificial intelligence has not yet been fully deployed in all industries. It is particularly difficult to develop technologies in areas that are suffering from liabilities, such as news media. The following section presents the nascent literature on artificial intelligence in journalism and discusses the potential that AI brings to the news industry.

2.2. Artificial Intelligence in the News Industry

Artificial intelligence has demonstrated potential in numerous experimental studies, particularly in scientific and technological fields. Although it is hard to estimate the cost of creating and implementing an artificial intelligence application without delving into one's project's details, many experts understand that the development of AI systems has reduced costs in recent years. However, AI still requires specialized expertise, which makes it difficult to compete with Silicon Valley companies. Big tech companies leverage their control over AI by actively acquiring startup companies that are deploying AI solutions in an attempt to concentrate power and thwart any competitors (Linden 2017a).

A recent report revealed that a major bottleneck for the development of AI in news media is the talent competition. This competition not only involves attracting talent but also retaining professionals in newsrooms, which offer lower salaries in comparison to the tech industry (Cook et al. 2021). This newsroom brain drain works against the adoption of technologies in the news industry (Broussard et al. 2019; Lokot and Diakopoulos 2016).

Even under those conditions, news outlets around the world are embracing AI solutions in their newsrooms. In recent years, newsrooms began to increasingly automate news stories (Linden 2017b). Although machine learning algorithms are used to a certain extent for some of these projects, many projects still rely on simple automation that fills in the blanks of template stories and does not produce stories built on prior data (Biswal and Gouda 2020).

Although many of these journalistic bots do not use machine learning or NLP models, they rely on a series of steps to take (planning) and understand when to carry out a certain step (scheduling) in order to publish messages. Some researchers have identified four different categories of news bots: "the inputs and sources of input data; the outputs produced by the news bots; the algorithms that guide how a news bot turns inputs into outputs; and the function or intent of the news bot" (Lokot and Diakopoulos 2016, p. 696). In Brazil, newsrooms mainly rely on Twitter bots that use AI models, particularly NLP, machine learning, and planning, scheduling, and optimization to process large volumes of data and interact on digital media platforms (DalBen and Jurno 2021). However, bots have also been used with malicious intent in several recent events, such as the 2016 US elections and the Brexit campaign (Bastos and Mercea 2018).

Many people expect AI techniques to reduce the costs of investigative journalism (Broussard 2015). However, AI models are generally built for a particular story, which means that these algorithms must be created and trained again for novel projects. As a result, high initial investments cannot be amortized over multiple products (Stray 2019). Similarly, investigative news projects that rely on computer vision require significant investments to build technological infrastructure and hire highly qualified personnel to develop such codes (de-Lima-Santos and Salaverría 2021). Furthermore, AI models are usually trained using old and biased datasets, which can generate many ethical complications (Guzman and Lewis 2020).

Traditional news outlets have successfully deployed AI projects in their newsrooms, such as *The New York Times* (NYT), *The Washington Post*, and *Associated Press* (Chan-Olmsted 2019). However, AI is expensive, even for them (Broussard et al. 2019). Furthermore, elite news organizations have to make a great deal of progress to realize what is possible with AI in their newsrooms. For example, the release of GPT-3 in 2020 evolved machine learning models for text to the next level. This "autoregressive language model with 175 billion parameters, $10\times$ more than any previous non-sparse language model" (Brown et al. 2020, p. 1), can perform many different tasks such as article generation, translation, summarization, and prediction and uses less computing power (Gage 2020). However, this applicability also brings risks such as the creation of distorted content, which can be used to deceive the public.

These advances in AI-related technologies have the potential to significantly disrupt the nature of human–machine interactions. AI technologies have proven to be in short-, medium-, and long-term part of a broader reconfiguration of the news industry, which started with digitalization and the creation of the internet (Broussard et al. 2019; Erdal 2011). AI is not a silver bullet for journalism, but it is a new tool that requires members of the news industry to possess more understanding to further support and bolster AI capabilities in newsrooms. AI-enforcement mechanisms are important for ensuring that AI systems adhere to legal and ethical guidelines without explicitly considering the power structures between various stakeholders (Broussard et al. 2019). For this reason, understanding the different subfields of AI is important.

3. Methods

To understand the evolution of AI in the news industry, one can turn to the list of case studies from JournalismAI. This is a global initiative led by the LSE journalism's think-tank, Polis, which aims to expand the knowledge about AI-powered technologies in newsrooms through collaborative projects. The institute offers a network for best practices and innovation sharing and also produces research reports and training materials for the news industry. On one of the fronts, JournalismAI built a database of cases that are at the intersection of AI and journalism. For this dataset, the team collected the best case studies that they could find and created a form for organizations and practitioners to submit examples of AI applications being used in the news industry. According to JournalismAI's portal, the database aims to "ensure that everyone in our global network can learn from and be inspired by these creative applications of AI in journalism" (JournalismAI n.d.).

The list comprises 102 items. We excluded nine entries that could not be accessed through the available URLs or that we could not find information about on the internet. Although we recognize that this database does not represent an extensive list, it should provide some idea of how AI is being developed in the news industry. To analyze the list, we built on the previous literature and classified these cases according to the types of AI applications that they use. As illustrated in Figure 2, the available artificial intelligence literature suggests that there are seven major areas of AI: (i) machine learning; (ii) natural language processing (NLP); (iii) speech recognition; (iv) expert systems; (v) planning, scheduling, and optimization; (vi) robotics; and (vii) computer vision. Based on these subfields, we formulated our categories of analysis and, from them, developed a theoretical discussion of our findings, highlighting examples that illuminate the application of AI in the news media industry.

Although we recognize that some subareas are present within each of these areas (for example, deep learning and predictive analytics are part of machine learning), our objective is to provide a macro approach to demonstrate how those topics work together in the journalistic field. We also acknowledge that certain projects rely on more than one subfield of AI. In this case, we classified only the two major subfields of AI found in these projects. Each author qualitatively analyzed each item from the database and coded it separately. During the process, we made annotations about each project to facilitate the discussion of discrepancies. After this process, we reunited and discussed the differences in classification. In total, 12 news products were initially classified differently.

Additionally, we analyzed where these projects were developed to obtain a better understanding of the distribution of AI-powered technologies in newsrooms worldwide. Our findings brought into perspective newsrooms' use of AI, especially regarding its use beyond the traditional forms of machine learning applications and natural language processing algorithms by detailing the most relevant aspects of these cases. In the following section, we present our findings based on our explorative research.
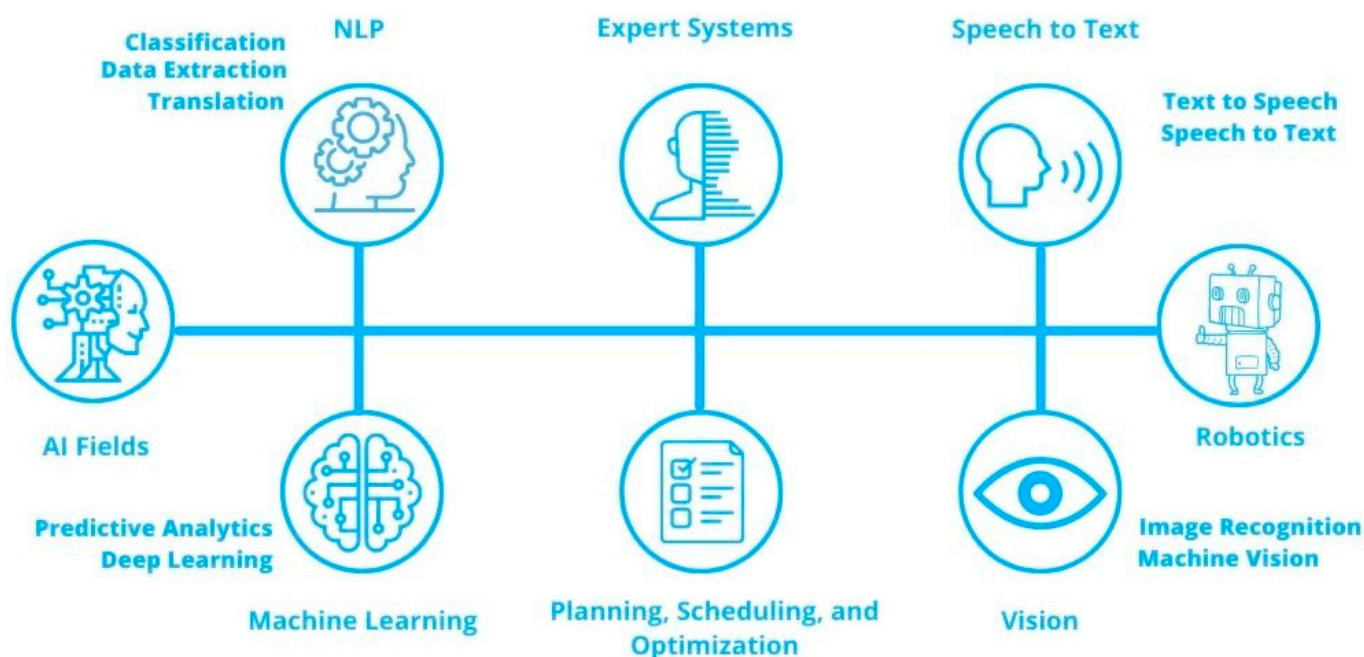
Figure 2. The seven fields of AI (Source: Authors).

4. Findings

4.1. An Overview of the AI in News Media

Overall, our findings reveal that most AI applications in the news industry are being developed in two regions: the Americas (43.01%, *N* = 40) and Europe (39.78%, *N* = 37). Our findings resemble those of prior studies on other technological developments in the news media ecosystem. For example, data journalism was first developed in the United States and European nations such as the United Kingdom and the Nordic countries (Appelgren and Nygren 2014; Borges-Rey 2016; Parasie and Dagiral 2013; Young and Hermida 2015). More recently scholars have identified developments in non-Western countries (Mutsvairo 2019). This is important because, as happened with data journalism, many have assumed the ubiquity and growing universal acceptance of AI in the news industry when this is not the reality. Artificial intelligence has been unequally developed since it is far too expensive for the majority of media companies. This limits most newsrooms' abilities to adopt cutting-edge technologies in the Global South (Linden 2017a). In our list, only 5.38% (*N* = 5) of the cases were from Asia, and 2.15% (*N* = 2) were from Oceania. The application of AI in African newsrooms appears to be very limited, as shown in prior studies (Kothari and Cruikshank 2021; Munoriyarwa et al. 2021).

Looking more specifically at the Americas, which also includes Latin America and Canada, only 4.3% (*N* = 4) of all the cases belong to this region. Although prior studies have highlighted some examples of Latin American organizations (de-Lima-Santos and Mesquita 2021b). The rest (38.71%, *N* = 6) belong to the US. Only 9.68% of these cases encompass global projects—that is, projects that reach wider audiences and different continents. Figure 3 depicts these subfields of AI in relation to each continent.
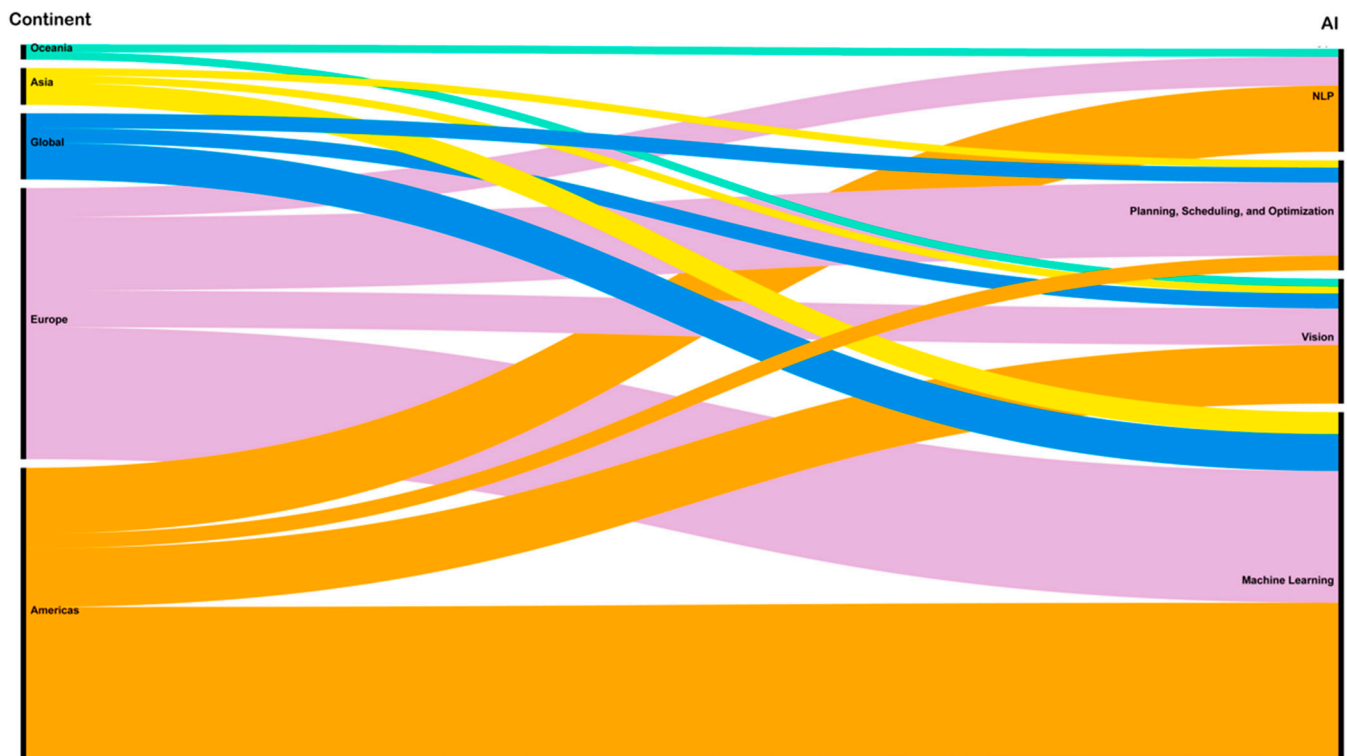
Figure 3. Fields of AI in relation to the continents (Source: Authors).

Regarding AI subfields, some researchers have noticed that the most common subfield is machine learning (66.67%, *N* = 62). This is followed by applications that rely on computer vision (18.28%, *N* = 17) and planning, scheduling, and optimization (16.13%, *N* = 15). NLP models learn a non-trivial amount of linguistic knowledge, which makes it harder for them to replicate it in different languages, such as Portuguese (Rodrigues et al. 2014). Although it can be found in computer science literature on fine-tuning models for the English language, NLP requires higher layers of specific knowledge for other languages, which impedes its consistent implementation in the news industry. For this reason, NLP models represent 15.05% (*N* = 14) of our sample.

However, the subfields of speech recognition, expert systems, and robotics are not present in our sample. Within such a fluid environment, increasing complexity is needed to deploy projects in these subfields of artificial intelligence in the news ecosystem since they require more specialized skills and since some of them require an innovative approach to develop novel products that were not previously available in the industry. This is happening at a moment when media companies are trying to survive uncertain times (Chan-Olmsted 2019). Such integration would require significant alignment within an organization and resource investments that news outlets do not foresee becoming obtainable in the near future (Locker et al. 2019).

However, we could note that most of these projects rely on external funds for deployment. Most funding comes from Google's Digital News Innovation (DNI) grants. Currently, news media is working with technology companies in a relationship that could be colloquially described as "frenemies" (Rashidian et al. 2018). On one hand, tech companies broke news media's business models; on the other hand, big tech platforms have become a major source of funding and a disruptor of innovation in the news industry (Rashidian 2020). Additionally, news organizations rely on third-party platforms such as Google's incubator Jigsaw to help them develop AI solutions for their business, echoing findings from previous studies (Cook et al. 2021). In the following subsections, we describe how the three main subfields of AI are being developed in the news industry.

4.2. Machine Learning and Its Applications in the Journalistic Field

Although many of these organizations rely on external funding to develop some projects, members of the news media are looking for ways to improve their strategies and boost their revenue streams. In our cases, two major trends appeared. First, the use of machine learning models was commonly found in projects that aimed to learn about readers' interests to boost engagement using a content recommendation engine. For example, *The New York Times*, *New Zealand Media and Entertainment* (NZME), and *Toutiao* (China) launched AI-powered news applications with this intent. In the same vein, *The Times* (UK) halved digital subscriber churn using tailored emails for its subscribers.

Second, a growing number of news outlets have built paywalls that bend to the individual reader or predict subscription cancellation. The Wall Street Journal is an example of this: the newspaper deployed a machine learning model that would allow non-subscribers to sample some stories. Using this machine learning algorithm, The Wall Street Journal could build a more flexible paywall that would inform news managers of the kinds of stories users are interested in. Similarly, third-party organizations began to build these solutions and offer them to newsrooms, such as Piano in the US and Deep BI in the UK; this echoes findings from previous studies about the emergence of new players in this field (Cook et al. 2021).

Although machine learning is the most popular subfield of AI since it helps computers improve their performance based on previous experiences (Russell and Norvig 2021), it is also the most common model because it can be applied to other AI subfields. One could easily identify some projects that have combined machine learning algorithms with NLP, computer vision, and planning, scheduling, and optimization. These subfields can be merged because these different components are entangled in the AI space (Castro and New 2016). Thus, the field of planning, scheduling, and optimization is one of the subfields of AI that is commonly applied in conjunction with machine learning.

4.3. Computer Vision to Investigative Reporting

Until recently, AI only worked at a limited capacity since technologists had to program a wide array of functions into a system to mimic human intelligence, which required massive computing power with low throughput. This has changed due to better hardware, more data, and better algorithms (Aronson 2018; Castro and New 2016; Hassaballah and Awad 2020; Whittaker 2019). Nonetheless, computer vision is still a subfield of AI that requires powerful tools to simulate human vision and enable a machine to learn to recognize abstract patterns in images (Castro and New 2016; Szeliski 2011).

Our sample included some small news outlets that have developed news stories using CV. The small news outlet Texty from Ukraine has relied on CV models to detect land that turned into lunar-looking landscapes due to illegal amber mining in the country. We also identified a major number of traditional news organizations that have adopted different approaches to the use of computer vision in their newsrooms. For example, *The New York Times* and Reuters have used CV algorithms in their news stories. NYT used computer vision algorithms to estimate 3D poses of sports athletes at live events, and Reuters used CV and satellite images to track urban expansion in the South China Sea. Reuter has also used this subfield of AI to enhance the search feature that is part of its video archive. The Chinese Xinhua News Agency has used CV in combination with other AI subfields to rebuild its newsroom by emphasizing human–machine collaboration to produce real-time stories.

In a similar vein, news organizations are using CV to detect fake images. In an event promoted by Google, Asian practitioners developed an app called Source, which is powered by Storyful. This app uses Google's AI technology to provide access to an image's public history, which can allow users to understand an image's provenance and any sort of manipulation that has been done to it.

4.4. Planning, Scheduling, and Optimization in News Media

Automated stories are commonly found in the subfield of planning, scheduling, and optimization. The process involves running data through an algorithm that organizes that data into a readable story. Therefore, using algorithms to plan, publish, and refine stories is usually implemented to produce data-driven stories; such stories often relate to crimes, natural disasters, elections, finance, and sports. After the success of the *LA Times'* QuakeBot, which can create a write-up within minutes of an earthquake, newsrooms started to embrace story automation (Salaverría and de-Lima-Santos 2020).

In general, narrative structures are repetitive, which makes them good candidates for automation (Carlson 2015; Dörr 2016; Graefe 2016; van Dalen 2012). For example, The Associated Press and *Newsday* automated the coverage of 124 school districts in the US, and *The Washington Post* published 850 automated articles in 2016. These examples highlight the potential that AI brings to news production since it allows newsrooms to produce more stories while using fewer human resources (Broussard et al. 2019). However, using automated stories raises ethical and quality concerns (Guzman and Lewis 2020).

Many news media outlets saw the COVID-19 outbreak as an opportunity to automate their production processes since the global death toll and infection rates were structured data that could fit into predictable story frames (Danzon-Chambaud 2021). For example, The Times (UK) automated a powerful charting tool to build graphics for coronavirus coverage. The BBC released a project called Salco ("Semi-Automated Local Content") to generate over 100 unique stories per month. The BBC allowed Salco's coverage to focus on local audiences who could learn about their hospitals' performance during the COVID-19 pandemic. These examples highlight another level of AI systems that includes the subfields of NLG and computer vision.

5. Discussion and Conclusions

Overall, this study argues that AI can take different forms in the news industry. Our findings reveal three major subfields that are more present in the news ecosystem: machine learning, computer vision, and planning, scheduling, and optimization. Machine learning is used in different parts of the news production workflow. However, we commonly found two applications in our cases. The first involves a great interest in boosting public engagement using machine learning recommendation engines. The second involves news outlets using machine learning models to adjust their business strategies to individual readers. For example, machine learning may be used to predict subscription cancellation or build paywalls that bend to the individual reader. Thus, machine learning algorithms are often used to strengthen news outlets' business models and boost revenue streams. In line with previous studies, third-party organizations build some of these solutions and sell them to newsrooms, such as Piano in the US and Deep BI in the UK (Cook et al. 2021). Similarly, large tech platforms, especially Google, provide solutions such as Jigsaw, which is a tool that is used to help community managers manage toxic comments or posts that might violate community guidelines (Rashidian 2020).

Automated journalism governs the subfield of planning, scheduling, and optimization. Although journalism is related to textual content, our findings suggest that NLP models are used less frequently in the industry than planning, scheduling, and optimization. We speculate that this might be associated with the fact that it is not easy to replicate NLP models in different languages, such as Portuguese (Rodrigues et al. 2014). Furthermore, automated journalism does not necessarily focus on adopting machine learning or NLP approaches. Automated journalism constitutes, at times, a basic application of computational models, which, in many cases, are used to fill in the blanks of template stories. These dumb algorithms—those which may adapt but ultimately seek to achieve very simple instructions—are far more likely to be found in the news industry, as they do not require much effort or time to deploy (Biswal and Gouda 2020).

Computer vision is a subfield of AI that helps practitioners deal with visual content in different ways. Thus far, most of the cases deal with CV as a tool for investigative reporting, including fact-checking content on social networks. According to this perspective, CV seems limited to being a one-goal project, which hampers its replication and, consequently, its application on a larger scale. Although we could identify small news outlets that rely on computer vision algorithms to produce investigative stories, we found that large newsrooms use these algorithms the most. We understand that since CV requires technological infrastructure, qualified personnel to develop codes, and a significant investment, only large newsrooms can afford it (de-Lima-Santos and Salaverría 2021). However, we found that most of the applications that use AI in the news industry rely on grants from big tech companies such as Google and Facebook to develop them (Rashidian et al. 2018; Rashidian 2020). This brings serious challenges to the development of technological innovations in news media since these organizations decide who receives their money, when they receive it, and where it goes.

We found that few of the news outlets we studied use social bots. Most of the bots that we researched were news bots that write stories. Social bots are the most easily applicable form of technological assistance in news production and dissemination (DalBen and Jurno 2021; Lokot and Diakopoulos 2016). However, we believe that there are two reasons why they did not appear on this list. First, social bots such as Twitter bots do not necessarily use AI. We speculate that this might be the reason why social media bots did not appear on the list of cases. Second, this list is not extensive enough, which might leave some examples of AI applications in the news industry out. This can result in a biased sample of the population. Because of this potential limitation, we treated the study as an initial approach to the subject. Future research could explore this topic, especially from practitioners' points of view. Furthermore, it would be interesting to have more ethnographic studies about the development of AI in newsrooms since they would shed light on limitations and hurdles that relate to deploying AI algorithms in the field. Given that AI has ethical considerations and societal consequences, these computational models need to be constructed and simulated with the most up-to-date datasets to reduce biased behaviors, which can affect the public and professionals involved (Guzman and Lewis 2020). Future studies can address these ethical and societal consequences in each subfield.

In conclusion, we were able to highlight the different uses of AI subfields in the news industry, despite the limitations discussed. Our study contributes to scholarly literature by stressing the limits and opportunities that relate to using AI in news media and providing input for practitioners to expand its applicability.

References

Appelgren, Ester, and Gunnar Nygren. 2014. Data Journalism in Sweden. *Digital Journalism* 2: 394–405. [CrossRef]

Aronson, Jay D. 2018. Computer Vision and Machine Learning for Human Rights Video Analysis: Case Studies, Possibilities, Concerns, and Limitations. *Law and Social Inquiry* 43: 1188–209. [CrossRef]

Bastos, Marco, and Dan Mercea. 2018. The Public Accountability of Social Platforms: Lessons from a Study on Bots and Trolls in the Brexit Campaign. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 376. [CrossRef]

Biswal, Santosh Kumar, and Nikhil Kumar Gouda. 2020. *Artificial Intelligence in Journalism: A Boon or Bane?* Singapore: Springer, pp. 155–67. [CrossRef]

Boczkowski, P. J. 2004. *Digitizing the News: Innovation in Online Newspapers. Inside Technology.* Cambridge and London: MIT Press, pp. 1–360.

Borges-Rey, Eddy. 2016. Unravelling Data Journalism: A Study of Data Journalism Practice in British Newsrooms. *Journalism Practice* 10: 833–43. [CrossRef]

Broussard, Meredith. 2015. Artificial Intelligence for Investigative Reporting: Using an Expert System to Enhance Journalists' Ability to Discover Original Public Affairs Stories. *Digital Journalism* 3: 814–31. [CrossRef]

Broussard, Meredith. 2018. *Artificial Unintelligence: How Computers Misunderstand the World*, 1st ed. Cambridge: MIT Press.

Broussard, Meredith, Nicholas Diakopoulos, Andrea L. Guzman, Rediet Abebe, Michel Dupagne, and Ching Hua Chuan. 2019. Artificial Intelligence and Journalism. *Journalism and Mass Communication Quarterly* 96: 673–95. [CrossRef]

Brown, Tom B., Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, and et al. 2020. Language Models Are Few-Shot Learners. *Advances in Neural Information Processing Systems*. Available online: https://arxiv.org/abs/2005.14165v4 (accessed on 1 November 2021).

Carlson, Matt. 2015. The Robotic Reporter: Automated Journalism and the Redefinition of Labor, Compositional Forms, and Journalistic Authority. *Digital Journalism* 3: 416–31. [CrossRef]

Castro, Daniel, and Joshua New. 2016. *The Promise of Artificial Intelligence. Washington, DC and Brussels.* Available online: https://www2.datainnovation.org/2016-promise-of-ai.pdf (accessed on 1 November 2021).

Chan-Olmsted, Sylvia M. 2019. A Review of Artificial Intelligence Adoptions in the Media Industry. *International Journal on Media Management* 21: 193–215. [CrossRef]

Coddington, Mark. 2015. Clarifying Journalism's Quantitative Turn. *Digital Journalism* 3: 331–48. [CrossRef]

Cook, Clare, Emiliana Garcia, Heghine Gyulnazaryan, Juan Melano, Jakub Parusinski, and Alex Sabadan. 2021. *The Next Wave of Disruption: Emerging Market Media Use of Artificial Intelligence and Machine Learning*, 1st ed. Edited by Robert Shaw. Copenhagen: International Media Support (IMS).

DalBen, Silvia, and Amanda Chevtchouk Jurno. 2021. More than Code: The Complex Network That Involves Journalism Production in Five Brazilian Robot Initiatives Cite This Paper Related Papers. *ISOJ* 11: 111–37.

Danzon-Chambaud, Samuel. 2021. Covering COVID-19 with Automated News. Columbia Journalism Review. Available online: https://www.cjr.org/tow_center_reports/covering-covid-automated-news.php (accessed on 6 August 2021).

de-Lima-Santos, Mathias-Felipe, and Lucia Mesquita. 2021a. Data Journalism Beyond Technological Determinism. *Journalism Studies* 22: 1416–35. [CrossRef]

de-Lima-Santos, Mathias-Felipe, and Lucia Mesquita. 2021b. A Challenging Future for the Latin American News Media Industry. In *Journalism, Data and Technology in Latin America*, 1st ed. Edited by Ramón Salaverría and Mathias-Felipe De-Lima-Santos. Cham: Palgrave Macmillan, pp. 229–62. [CrossRef]

de-Lima-Santos, Mathias-Felipe, and Ramón Salaverría. 2021. From Data Journalism to Artificial Intelligence: Challenges Faced by La Nación in Implementing Computer Vision in News Reporting. *Palabra Clave* 24: 1–40. [CrossRef]

Deloitte. 2014. *Demystifying Artificial Intelligence*. New York: Deloitte.

Diakopoulos, Nicholas. 2020. Computational News Discovery: Towards Design Considerations for Editorial Orientation Algorithms in Journalism. *Digital Journalism* 8: 945–67. [CrossRef]

Dörr, Konstantin Nicholas. 2016. Mapping the Field of Algorithmic Journalism. *Digital Journalism* 4: 700–22. [CrossRef]

Erdal, Ivar John. 2011. Coming to Terms with Convergence Journalism: Cross-Media as a Theoretical and Analytical Concept. *Convergence: The International Journal of Research into New Media Technologies* 17: 213–23. [CrossRef]

Ferrer-Conill, Raul, and Edson C. Tandoc. 2018. The Audience-Oriented Editor: Making Sense of the Audience in the Newsroom. *Digital Journalism* 6: 436–53. [CrossRef]

Future Today Institute. 2018. 2019 Industry Trends: Journalism, Media, Technology. Available online: https://futuretodayinstitute.com/ (accessed on 30 September 2021).

Gage, Justin. 2020. *What's GPT-3?—Technically. Technically.* Available online: https://technically.substack.com/p/whats-gpt-3 (accessed on 20 July 2020).

Gelgel, Ni Made Ras Amanda. 2020. Will Technology Take over Journalism? *Informasi* 50: 5–10. [CrossRef]

Graefe, Andreas. 2016. *Guide to Automated Journalism*. New York: Tow Center for Digital Journalism, Columbia University. [CrossRef]

Guzman, Andrea L., and Seth C. Lewis. 2020. Artificial Intelligence and Communication: A Human–Machine Communication Research Agenda. *New Media & Society* 22: 70–86. [CrossRef]

Harvard, Jonas. 2020. Post-Hype Uses of Drones in News Reporting: Revealing the Site and Presenting Scope. *Media and Communication* 8: 85–92. [CrossRef]

Hassaballah, Mahmoud, and Ali Ismail Awad. 2020. *Deep Learning in Computer Vision*, 1st ed. Edited by Mahmoud Hassaballah and Ali Ismail Awad. Boca Raton: CRC Press/Taylor and Francis. [CrossRef]

Hermida, Alfred, and Mary Lynn Young. 2019. *Data Journalism and the Regeneration of News*, 1st ed. London: Routledge. [CrossRef]

Hernandez Serrano, Maria Jose, Anita Greenhill, and Gary Graham. 2015. Transforming the News Value Chain in the Social Era: A Community Perspective. Edited by Dr. Gary Graham. *Supply Chain Management: An International Journal* 20: 313–26. [CrossRef]

Jamil, Sadia. 2020. Artificial Intelligence and Journalistic Practice: The Crossroads of Obstacles and Opportunities for the Pakistani Journalists. *Journalism Practice* 1–23. [CrossRef]

JournalismAI. n.d. Case Studies. *JournalismAI*. Available online: https://www.lse.ac.uk/media-and-communications/polis/JournalismAI/Case-studies (accessed on 30 September 2021).

Kang, Seok, Erin O'Brien, Arturo Villarreal, Wansoo Lee, and Chad Mahood. 2019. Immersive Journalism and Telepresence. *Digital Journalism* 7: 294–313. [CrossRef]

Kothari, Ammina, and Sally Ann Cruikshank. 2021. Artificial Intelligence and Journalism: An Agenda for Journalism Research in Africa. *African Journalism Studies* 1–17. [CrossRef]

Krumsvik, Arne H, Stefania Milan, Niamh Ní Bhroin, and Tanja Storsul. 2019. Making (Sense of) Media Innovations. In *Making Media*, 1st ed. Edited by Mark Deuze and Mirjam Prenger. Amsterdam: Amsterdam University Press, pp. 193–206. [CrossRef]

Lewis, Seth C., Andrea L. Guzman, and Thomas R. Schmidt. 2019. Automation, Journalism, and Human–Machine Communication: Rethinking Roles and Relationships of Humans and Machines in News. *Digital Journalism* 7: 409–27. [CrossRef]

Linden, Carl-Gustav. 2017a. Algorithms for Journalism: The Future of News Work. *The Journal of Media Innovations* 4: 60–76. [CrossRef]

Linden, Carl-Gustav. 2017b. Decades of Automation in the Newsroom. *Digital Journalism* 5: 123–40. [CrossRef]

Locker, Mic, Jeff Loucks, Susanne Hupfer, and David Jarvis. 2019. *Seasoned Explorers: How Experienced TMT Organizations Are Navigating AI*. New York: Deloitte.

Lokot, Tetyana, and Nicholas Diakopoulos. 2016. News Bots: Automating News and Information Dissemination on Twitter. *Digital Journalism* 4: 682–99. [CrossRef]

Marr, D. 2010. *Vision: A Computational Investigation Into the Human Representation and Processing of Visual Information*. Cambridge: The MIT Press.

McCarthy, John. 1998. What Is Artificial Intelligence? *CogPrints*; Stanford. Available online: http://cogprints.org/412/2/whatisai.ps (accessed on 30 September 2021).

Munoriyarwa, Allen, Sarah Chiumbu, and Gilbert Motsaathebe. 2021. Artificial Intelligence Practices in Everyday News Production: The Case of South Africa's Mainstream Newsrooms. *Journalism Practice Online First*. [CrossRef]

Mutsvairo, Bruce. 2019. Challenges Facing Development of Data Journalism in Non-Western Societies. *Digital Journalism* 7: 1289–94. [CrossRef]

Nelson, Jacob L., and Edson C. Tandoc. 2018. Doing 'Well' or Doing 'Good': What Audience Analytics Reveal About Journalism's Competing Goals. *Journalism Studies* 20: 13. [CrossRef]

Örnebring, Henrik. 2010. Technology and Journalism-as-Labour: Historical Perspectives. *Journalism* 11: 57–74. [CrossRef]

Ortiz Freuler, J., and C. Iglesias. 2018. *Algorithms e Inteligencia Artificial En Latin America: Un Estudio de Implementaciones Por Parte de Gobiernos En Argentina y Uruguay*. Available online: www.webfoundation.org (accessed on 1 November 2021).

Parasie, Sylvain, and Eric Dagiral. 2013. Data-Driven Journalism and the Public Good: 'Computer-Assisted-Reporters' and 'Programmer-Journalists' in Chicago. *New Media & Society* 15: 853–71. [CrossRef]

Paulussen, Steve. 2016. Innovation in the Newsroom. In *The SAGE Handbook of Digital Journalism*, 1st ed. Edited by Tamara Witschge, C. W. Anderson, David Domingo and Alfred Hermida. London: SAGE Publications Ltd., pp. 192–206. [CrossRef]

Rashidian, Nushin. 2020. *Platforms and Publishers: The Great Pandemic Funding Push. Columbia Journalism Review*. Available online: https://www.cjr.org/tow_center_reports/platforms-publishers-pandemic-funding-news.php (accessed on 17 December 2020).

Rashidian, Nushin, Pete Brown, Elizabeth Hansen, Emily Bell, Jonathan Albright, Abigail Hartstone, Elizabeth Hansen-With, Emily Bell, Jonathan Albright, and Abigail Hartstone. 2018. Friend and Foe: The Platform Press at the Heart of Journalism. *Columbia Journalism Review*. [CrossRef]

Rodrigues, Ricardo, Hugo Gonçalo Oliveira, and Paulo Gomes. 2014. Lemport: A High-Accuracy Cross-Platform Lemmatizer for Portuguese. *OpenAccess Series in Informatics* 38: 267–74. [CrossRef]

Russell, Stuart, and Peter Norvig, eds. 2021. *Artificial Intelligence: A Modern Approach*, 4th ed. Harlow, UK: Pearson, Available online: https://www.pearson.com/us/higher-education/program/Russell-Artificial-Intelligence-A-Modern-Approach-4th-Edition/PGM1263338.html (accessed on 1 November 2021).

Salaverría, Ramón, and Mathias-Felipe de-Lima-Santos. 2020. Towards Ubiquitous Journalism: Impacts of IoT on News. In *Journalistic Metamorphosis: Media Transformation in the Digital Age Volume 70 of Studies in Big Data*, 1st ed. Edited by Jorge Vázquez-Herrero, Sabela Direito-Rebollal, Alba Silva-Rodríguez and Xosé López-García. Cham: Springer International Publishing, pp. 1–15. [CrossRef]

Stray, Jonathan. 2019. Making Artificial Intelligence Work for Investigative Journalism. *Digital Journalism* 7: 1–22. [CrossRef]

Szeliski, Richard. 2011. *Computer Vision*. Texts in Computer Science. London: Springer. [CrossRef]

van Dalen, Arjen. 2012. The Algorithms behind the Headlines: How Machine-Written News Redefines the Core Skills of Human Journalists. *Journalism Practice* 6: 648–58. [CrossRef]

Whittaker, Jason. 2019. *Tech Giants, Artificial Intelligence, and the Future of Journalism*. London: Routledge. [CrossRef]

Young, Mary Lynn, and Alfred Hermida. 2015. From Mr. and Mrs. Outlier To Central Tendencies. *Digital Journalism* 3: 381–97. [CrossRef]