Journal of Information Technology Management

Print ISSN: 2008-5893

Online ISSN: 2423-5059

Development of Robot Journalism Application: Tweets of News Content in the Turkish Language Shared by a Bot

Hikmet Tosyalı *

*Corresponding Author, Assistant Professor, Faculty of Communication, Maltepe University, Istanbul, Turkey. E-mail: hikmettosyali@maltepe.edu.tr

Çiğdem Aytekin

Associate Professor, Faculty of Communication, Marmara University, Istanbul, Turkey. E-mail: cigdem.aytekin@marmara.edu.tr

Abstract

Today, news texts can be created automatically and presented to readers without human participation through technologies and methods such as big data, deep learning, and natural language generation. With this research, we have developed an application that can contribute to the literature regarding The Studies on Robot Journalism Applications with a technology-reductionist perspective. Robot journalism application named Robottan Al Haberi (the English equivalent of the application name is "get the news from the robot") produces news text by placing weather, exchange rates, and earthquake data in certain templates. The news texts, which are produced by placing the data in appropriate spaces on the template and with a maximum length of 280 characters, are automatically shared via the Twitter account @robottanalhaber. The weather information is shared once a day, the exchange rate information is shared three times a day, and the earthquake information is shared instantly. Here, we aim to produce automatic and short news by using the available structured data by placing them in specific news templates suggesting different options or a combination of them for different situations.

Keywords: Natural language generation, Artificial intelligence, Robot journalism, Data journalism, Newspapers.

DOI: 10.22059/jitm.2020.79335

Introduction

News refers to the information conveyed through the media and its organs regarding various events taking place all over the world. A journalist undertakes the task of delivering the news to readers. Today, news texts can be created automatically and presented to readers without human participation through technologies and methods such as big data, deep learning, and natural language generation (NLG). This led to the articulation of journalism with the concept of robot. Robot journalism practices are based on NLG. It can be argued that NLG has potential use in journalism, although it is used in other fields today as in the novel, named The Day a Computer Wrote a Novel, written jointly by people and artificial intelligence (Birer, 2016).

The researches on robot journalism are categorized according to their content in the literature section of this study. In the current study, we have developed an application that can contribute to the literature regarding the studies on robot journalism applications with a technology-reductionist perspective. Robot journalism application named *Robottan Al Haberi* produces news text by placing weather, exchange rates, and earthquake data in certain templates. The current study is important as it is the first academic study in Turkey to develop robot journalism practice. We aim to produce automatic and short news by using the available structured data by placing them in specific news templates suggesting different options or a combination of them for different situations. We show how this news is automatically shared on a Twitter account (i.e., @robottanalhaber). In the study, we expected that the account would be of interest and followed by users due to the transfer of important information on the three issues mentioned above. Finally, evaluations regarding the account with these features were discussed.

Robot Journalism

As Tornoe (2014) indicated, smart machines have prepared us for an inevitable future where they are trying to capture and enslave humankind. It is possible to see the reflections of this situation in many areas today. Now, robots have already entered the field of journalism. While artificial intelligence has begun to occupy an area where the human factor has traditionally dominated, especially, it affects journalism as a profession that has passed through a turbulent period. (Túñez-López et al., 2018). The technological advances and the ways to deliver a new content change the way journalism works and encourages new forms and differences. Many of these changes focus on the availability of online tools and data. Data journalism, geo-journalism, computer-aided reporting, and robot journalism are some of the new ways, and the role and function of the journalist have been changing (Szews, 2018).

The editor of Physics Today Journal, Charles Day (2018), stated in his article, published in Computing in Science & Engineering that he first learned about the robot journalism five

years ago: “A start-up called StatSheet had created software that took electronic versions of baseball scorecards and turned them into news reports. Known since 2011 as Automated Insights, the company and its competitors work with media outlets to automatically generate thousands of news stories a year. For now, the stories are confined to game summaries, earnings reports, and other topics in which the principal content resides in readily harvestable data. But there are signs that robot writers are expanding their reach and becoming more capable.” (p. 101)

Without a doubt, algorithms today can produce real and accurate content to attract users. Artificial intelligence and automation provide robust clues for a future where robots and software will reduce the need for human labor. Therefore, journalism has become a new area where robots are used to produce content (Shekhar, 2016).

According to a definition, robot journalism is the use of software algorithms designed to produce written articles. The algorithm is mainly structured to produce a news template. The sentences, used by a journalist while writing an article, are utilized in creating a template. The algorithms perform journalists’ routine writing tasks, that can be formulated and are mostly based on quantitative data, with higher speed and accuracy. Thus, news centers can distribute these automatically generated articles to the channels in real-time. When software is implemented for the first time, all articles are reviewed by people first. If the software is successful, the articles are automatically published after a final check (McCartney, 2015). In addition to the definition above, Latar (2018) stated that robot journalism has a second function: Software development to automatically extract new information from large data stacks. Therefore, the concept of robot journalism is significantly related to big data journalism since data are essential causes of robot journalism (Narin, 2017).

On the one hand, different terms are used, such as automated journalism, news writing boots, algorithmic journalism, computational journalism, bot journalism, Robo journalism, and web 3.0 journalism. On the other hand, according to Loosen (2018), the news cycle broadens from the inner circle involving observation, production, distribution, and consumption to the outer circle involving automated production, algorithmic data processing, and data generation. Thus, Loosen defines automated journalism and algorithmic journalism as data journalism.

Vállez and Codina (2018) suggested that the computational journalism allows the transformation of disciplines to evolve into journalism and to adapt to new environments, and it is characterized by interdisciplinarity and multidimensionality, which gives it a dynamic character. It is characterized by intensive use of technology that not only for the news creation process but also as an aid for news writing and ease of personalizing content for different environments. The computerized journalism focuses on answering the following questions:

How are new stories discovered using data and algorithms? How are visually engaging stories understood? How to personalize content for different types of readers?

In this study, we will use the term robot journalism and follow the definition of McCartney (2015). We will consider the robot journalism as algorithms that automatically convert the data into the news texts without human participation. Examining the existing literature on robot journalism, which is in a limited number, studies have focused on the following topics: Perceptions on how robots write the news (e.g., Wölker & Powell, 2018); optimistic (e.g., Tornoe, 2014), semi-optimistic (e.g., Dörr, 2016), and pessimistic (e.g., Bucher, 2017) attitudes towards robot journalism; ethical and legal problems regarding robot journalism (e.g., McCartney, 2015); application development (e.g., Sim & Shin, 2016). We reviewed these topics, respectively.

Perceptions about News Written by Robots

Clerwall (2014) investigated how readers would interpret the software-generated content against similar content written by a journalist. In the study, 30 participants were exposed to different news articles written by a journalist or produced from software. Then, they answered questions about how they perceived the article (e.g., overall quality, reliability, impartiality). Findings suggested that although the participants perceived the content created by the software as descriptive and boring, they could not distinguish it from the content written by journalists. In addition, the contents created by journalists were perceived as objective. The researcher stated that the bots could not replace the real journalists. According to Clerwall, bots would be given tasks to write ordinary stories based on data; however, more complex tasks such as narration and eyewitness reporting would be left to human journalists.

Kaa and Krahmer (2014) focused on the perceived reliability of news stories written by a robot, which investigates the differences and similarities between journalists and newsreaders. There were 232 participants, and 64 of them were journalists. The participants were asked to evaluate the perceived expertise and reliability levels of four news articles written in Dutch. Sports or finance topics were chosen as news stories. A robot or human journalist was chosen as the author to determine the perceived reliability of the news writer and the content of the news story. There was no significant difference in news readers' perceptions regarding the level of reliability and expertise of the robot writer and the journalist. However, a significant difference was found on the perceived reliability of the news source within the group of journalists. Journalists reported that the reliability of the human journalist was higher than that of a robot. Moreover, journalists stated that the perceived expertise of the robot was greater.

In another study, a group of researchers discussed the implications of the news written by a robot in journalism. In the study, the perceptions of the public and the journalists about an article written by a robot were compared with the article of a human journalist.

Furthermore, the effect of a manipulation created through the declaration of the author (a human author or a robot author) was explored through a series of experiments. Regarding the public's assessment of the quality of both articles, there was an interaction effect suggesting that the public gave higher scores to the article written by the robot when the author was declared as a robot. However, the robot article was given lower scores when the author was declared as a human journalist through the manipulation. This result is an indication of the public's negative attitude towards journalists. Hypothetically, the journalists were expected to give higher scores for a journalist's article and lower scores for an article written by the robot. In contrast to the expectation of the researchers, the journalists gave higher scores when the author was declared as a robot, but when the author was declared as a human journalist, they gave lower scores to the quality of the article written by a robot in reality. (Jung et al., 2017).

A study conducted by Sarılar (2019) examined whether there would be any difference between the robot journalist and the human journalist in terms of the evaluations on an article by a group of participants. The participants were read two articles about the earthquake, one written by a human journalist and the other by a robot journalist, and the results were compared. Most of the group correctly knew which article was written by the robot. Likewise, most of the group stated that the article written by the robot was more objective and understandable. The participants described the article written by the robot as more fun because it did not tire the person while reading it. However, articles written by the robot were evaluated as weaker in conveying emotions to the reader, and the participants reported that the storytelling was not realized and it was not successful in descriptions compared to the other article written by a human journalist.

Another empirical study conducted by Wölker and Powell (2018) aimed to investigate how European newsreaders (N = 300) would perceive three forms of journalism (i.e., the human, the robot, and combined journalism created by both the human and the robot) in terms of content and source reliability. The findings demonstrated that reliability perceptions regarding the contents and the sources did not differ between robot journalism and the combined journalism, including both the human and the robot. Only the news created by the robot in the field of sports was perceived as more reliable than the human journalist. As a result, it was shown that the effects of robots on journalism quality were highly indistinguishable for European readers.

Robots have increasingly accepted in the US and Chinese newsrooms, but more research is needed on the extent to which they are accepted by news users. Based on this need, a cross-cultural study conducted by Zheng et al. (2018) aimed to comparatively examine automated news perceptions of the users in terms of how they would perceive the quality of the news created by the robots, how much they would love this news, and how much they would trust the news. The results showed that US and Chinese users had similar

perceptions about the news generated by the robots. Users did not perceive automated content linearly but evaluated it by taking into account the interaction of authors (human journalists or robots), media outlets (traditional or new media), and cultural background (US or Chinese users).

The automated news generation can be easily done with NLG methods in areas where the data capacity is huge, and the data structure is known. Melin et al. (2018) developed an automation system for automated news generation, which they called Valtteri. Valtteri produced news articles about the 2017 Finnish municipal elections. Opinions were received from 152 evaluators to assess the quality of the articles and to identify aspects that need improvement. The articles were evaluated in the context of reliability, liking, quality, and representativeness. The evaluators reported that 21% of automated articles would have been written by journalists, and 10% of articles written by journalists would have been written by Valtteri. There was a significant positive correlation between the percentage of users making these inaccurate evaluations and the age of the evaluators. In general, men made fewer errors than women while identifying whether the author was a robot or a human journalist.

As seen, sometimes the objectivity dimension comes to the fore in the perception of the news written by the robots and it is articulated with the reliability dimension, but sometimes the tediousness of the content produced is concerned. Nevertheless, the fact that news written by robots is often indistinguishable from what the human journalist writes is important regarding the deterministic approach to robot journalism.

Optimistic, Semi-Optimistic, and Pessimistic Attitudes toward The Robot Journalism

Today, the robot journalism may be more common than most people would imagine. For example, artificial intelligence can scan articles that were searched to write this article. But can it compete with people analyzing the scanned content? This is an issue that has not been resolved yet. (Mayes, 2014). Tornoe (2014) had an optimistic attitude towards this issue, suggesting that since our childhood, smart machines have done an excellent job preparing us for an inevitable future that seeks to capture and enslave humans. In addition, he continued as follows: “Even Marshall Brain, founder of How Stuff Works and author of Robotic Nation, put journalists on notice as being one of nine professions that will ultimately be replaced by computers.” (p. 24)

In another example, while human journalists focus on larger projects in the Associated Press, robot journalists are responsible for writing routine business and sports stories. The robot journalism can be a practical and strategic step in writing complex stories for Associated Press and other publishers. According to Associated Press Strategy and Institutional Development Senior Vice President Jim Kennedy, automated story writing in

business and sports has increased significantly using robot journalism technology. (Young, 2016).

Ali and Hassoun (2019) stated that artificial intelligence technologies develop the work of journalists instead of changing the current understanding of journalism. Therefore, artificial intelligence does not pose a threat to professional journalism. Shekhar (2016) supported this idea by recommending not to worry about the rise of the robot journalism because the robot journalism has a role in creating news that requires only structured and quantitative data, not in mainstream journalism structured with human emotions or aspects. Thus, we can think of robots as an assistant reporter who processes all the tedious data numbers and turns them into readable content. The early signs of the skills of the robot reporters are considered promising. (Rutkin, 2014).

On the one hand, it is ensured that bots write news about sports, finance, and politics through artificial intelligence algorithms. On the other hand, some people still mistakenly think that the issue of text creation by bots is complicated. However, artificial intelligence algorithms are used to summarize scientific research articles and automatically create press releases and news stories. (Tatalovic, 2018).

Waddell (2019) provides journalists and editors with possible guidance on citation practices that can reduce the perceived bias and increase the perceived reliability of news created by the automation. Those who oppose the automation seemed to prefer automation to include references that define the collaboration between journalists and the algorithm. They also prefer to see references to clearly describe the role of automation from the beginning of an article. Therefore, while news media continues to face difficulties in communicating with politically polarized audiences, automation appears promising if it reminds journalists and readers that people are still involved in the news production process.

A study, conducted by Túñez-López et al. (2018) with the participation of 366 journalists, suggested that there is a lack of deep understanding of how artificial intelligence can directly affect journalism as a profession. None of the journalists participating in the survey mentioned the need to rethink their role in the process of creating the news.

Kim and Kim (2017) aimed to reveal the views of newspaper managers regarding what they would think about the placement of the robots in newsrooms. The data were collected from 24 different newspaper managers. According to the results, the business performance that arises with robot journalism and readers' willingness to read the news written by the robots are among the most important criteria for the managers to decide whether to integrate robot journalism into their business processes. On the other hand, the attitude of journalists towards robot journalism was evaluated in terms of business performance and changes in the

foreign market environment. When robot journalists are adopted, the decrease in the number of human journalists will most likely be chosen by the managers as an employment strategy.

Karlsen and Stavelin (2014) studied computational journalism as a craft applied in Norwegian newsrooms. Computational journalism is a continuation of traditional journalism, according to in-depth interviews with six expert practitioners working in the largest newsrooms in Norway. According to them, the benefit of computational journalism is not to save them from standard level work taking their time; the benefit is the development of new forms of data-driven and user-driven journalism with the potential to fulfill the traditional hopes and promises of journalism.

Traditional news production has been changing with the development of software that automatically generates the text in natural language from structured data. In his study, Dörr (2016) first addressed the question of whether generating text in natural language could fulfill professional journalistic functions at a technical level and claimed the existence of technological limitations in NLG. Secondly, while focusing on the economic potential of NLG in journalism, he expected its institutionalization at an organizational level. Based on these points of view, 13 semi-structured interviews were held with the representatives of the different service providers. The interviews reflected the current market situation in detail. (see more in Dörr, 2016). As a result, although the expansion of the NLG market is still at an early stage in journalism and the service providers and the products are scarce, NLG could perform the tasks of professional journalism at a technical level.

Narin (2017) examined the opportunities and the threats created by robot journalism, the change in the journalists' self-abilities, and whether new ethical problems were added to existing ones. According to her, the most striking effect of robot journalism applications would be unemployment that journalists might face. However, there is no such risk of unemployment for journalists who create different contents such as interviews, news analysis, and articles where human emotions, thoughts, and interactions are salient. On the other hand, with the convergence in computer sciences and communication sciences, robot journalism practices have started to become widespread. Due to this spread, the need for journalists with software and coding knowledge has increased. The prevalence of robot journalism brings about some ethical problems. Ethical problems caused by automated news written by algorithms could be listed as follows: dependence on public authorities from which data used in automated content is received, beliefs that automated content would be impartial, filtering, privacy, confidentiality, and security.

Dalen (2012) focused on how journalists force themselves to review their skills in terms of robot journalism. Based on this point, the researcher analyzed the reactions to the start of a sports website written by a machine. The journalists who evaluated the automatized news

content emphasized the ability of the machine to write complex sentences as an important skill defining the journalist and considered robot journalism as an opportunity. Moreover, contrary to previous studies that define journalists as conservative and defensive, human journalists who write about robot journalists do not reject new developments. There may be different reasons for this. For example, the journalists who have written about the impact of technological developments on journalism could be seen as early adopters that are more open to change in the journalistic community in general or the journalists writing about the automated content analysis can focus on a trend that does not directly affect their work. However, they can be more defensive when it comes to a direct impact on their work.

Another study was conducted with the participation of three national and six local media workers (Güz & Yeğen, 2018). Examining the responses of the participants via in-depth interview on the new media and robot journalism, although the new media affects traditional journalism positively, just like many other fields, in general, it was concluded that the professionals in the field are negatively affected by the process. Supporting this finding, Karaduman (2017) suggested that the journalism is a phenomenon whose subject is the human, therefore, even flawless robot journalism will always be missing one part. Similarly, Marconi and Siegman (2017) indicated that artificial intelligence could help expand journalism, but it will never replace journalism. For example, artificial intelligence can help the reporting process; however, journalists always need to put the pieces together, and they generate creative narratives.

Bucher (2017) attempted to reveal the link between computing and journalism through interviews with executive staff, editors, and developers at Scandinavian news organizations. There were three claims made by the researcher: 1) The machines have no instincts, 2) Democracy can never be personalized, and 3) The computation is something to consider rather than simulate. Thus, it was claimed that what cannot be calculated is not only a technical but also a deep social, cultural, political, and economic problem.

On the other hand, İrvan (2017) stated that a robot that started working in a newspaper in China wrote the first news, and this robot managed to write 300 characters news in a second about a local spring festival event. Robots have a stronger data analysis capacity than humans, and they can write news faster. Notwithstanding, robots could not replace human journalists because they are unable to interview and ask questions [yet].

As can be seen, some authors have an optimistic attitude that the robot journalists will replace the human journalists, while others have a semi-optimistic attitude by highlighting the function of the robot journalists to assist the human journalists. For example, robot journalists can play a supporting role in creating news that requires structured and quantitative data, or the algorithms can be implemented under the control of journalists and become user driven.

These are opportunities for journalists; however, there are some pessimistic concerns asking the following questions: What will be the business performance of robot journalism? Are readers eager to read the news written by a robot? How can social, cultural, and political processes be integrated into automated production?

Ethical and Legal Issues Related to The Robot Journalism

Many journalists consider automation as an opportunity to be more creative. Routine and tedious tasks can be transferred to bots, allowing journalists to focus more on creative tasks. However, the robot journalism raises some ethical issues that people should discuss such as data source authority, data disclosure, and authorship. (McCartney, 2015).

The reliance on algorithms for robot journalism has significant practical, socio-political, and psychological dimensions, and it affects news organizations, journalists, and readers both legally and professionally. Therefore, algorithmic authoring may be one of the most controversial but undiscovered aspects. In a study conducted by Montal and Reich (2017), a theoretical framework of algorithmic creativity in an interdisciplinary perspective was revealed by a field study. In this context, the citation issue, full information policies, and legal opinions regarding the studies produced by the algorithm were discussed. The study found that there were significant differences between the perception of authorship, typical citation regimes, and scientific literature. The results provided a consistent and comprehensive policy that supports the public interest in automated news to reduce these inconsistencies. For example, the recommended association policy for algorithmic content generation is as follows: If the line belongs to a single in-company programmer, it must be attributed to the software vendor or programmer; the algorithmic nature of the content should be clearly stated with full explanation (describing the software vendor or programmer's role in the organization); the data sources and algorithm methodology of a particular story should be detailed.

While robot journalism creates many new considerable opportunities, it also raises critical legal problems. For example, it is not only necessary to identify clear questions about what the algorithm should discuss in the article, but also how it overlaps with the rules of the traditional publishing, editorial control, privacy, and data protection framework. (Ombelet et al., 2016).

With the emergence of the news created by the software, specific ethical, moral, and operational issues must also be considered. Therefore, the political economy of the algorithm systems should be evaluated separately by publishers, advertisers, data producers, governments, and users. Furthermore, algorithms can be manipulated by humans, even in the most challenging print attempts. For this reason, journalism and media researchers should focus on the algorithm problem. (Lindén, 2017).

Implementing criminal and legal responsibility for abuse and neglect in the use of automated journalism may be the best tool to ensure the legitimacy of automated journalism. This seems to be the easiest way to respect ethical principles. In this perspective, some passing laws may be required to impose responsibility on the programmers. (Monti, 2018).

As can be seen, intellectual rights, transparency in content generation, channels used in obtaining data, data protection, the bias in algorithms, and legal responsibilities are some of the ethical and legal problems we could face in robot journalism. Therefore, essential dilemmas for robot journalism should be investigated in a multidimensional way through further studies, and legal regulations should be determined and added to the field of journalism in this direction.

The Studies on Robot Journalism Applications

Sim and Shin (2016) introduced an application that automatically writes articles called *Stock Robot* in their study. The necessary steps in Stock Robot include four phases: data collection and storage, critical event extraction, article content production, and article production. The robot collects and analyzes data on social issues and stock indices in the past two years. Twitter is used as a platform in this study. The produced article was uploaded to Twitter on an account without any errors.

Wordsmith is a platform created by Automated Insights. The platform collects vast quantities of measurable data from vertical fields dominated by detailed numerical data such as financial reports, real estate, and sports. It uses specific algorithms to compare that data. The platform then compiles them all in one story. It also uses preset parameters to create adjectives and punctuation marks (Tornoe, 2014).

Los Angeles Times has switched to automatic reporting on earthquakes. Quakebot, an algorithm created by database maker Ken Schwencke, creates a text on earthquakes with magnitudes greater than 3.0, with data from the U.S. Geological Research Unit. The algorithm adds a map to the text and presents it by creating a title within minutes of the earthquake's first detection. According to Schwencke, Quakebot helps journalists by eliminating most of the regular work in earthquake reporting (Tornoe, 2014). Rutkin (2014) says that the Los Angeles Times newspaper automatically publishes the news containing all the crucial details about the earthquake 3 minutes after the earthquake in California, and highlights this point: “The last sentence was the most important, it says, ‘this post was generated by an algorithm created by the author.’ This sentence got more attention”. Quakebot works by filling the gaps in the relevant places. When the U.S. Geological Survey Unit sends an e-mail alert about an earthquake, Schwenke’s bot parses the email, places the data in a template, and uploads the article to the newspaper’s content management system. It even sends an email to editors to see

notifications. In addition, the newspaper has other bots to compare neighborhoods for a project; in this way, articles on city murders are automatically generated on the website.

In their study, Lee et al. (2017) presented a prototype of an algorithm (Custombot's Algorithm) that creates personalized news articles on *information technologies and technology* for a particular theme, criterion, or item. When a personal preference is made, the algorithm analyzes the data, originates the most appropriate topic preferred by a person, and finally produces a news article on this topic. The Custombot Algorithm, which processes and analyzes data with inductive reasoning, considers the concepts of news angle and filter bubble. Text segmentation and private labeling are two of the tasks used in the system that allows creating custom labels and adding matching information to the appropriate places. Content, structure, and layout will be different for each article, depending on the category the user chooses. As a result, the article can be served as a new service provided by news agencies to meet the needs of each news consumer.

Verhulst et al. (2017) introduced Marvin, a futuristic newsroom simulation, in their study. The simulation consists of three scenarios involving regular journalistic tasks: Data analysis, writing, and interview. A total of nine participants with journalistic experience tested one or two of these scenarios in a special room representing the futuristic newsroom at the University of Amsterdam. As a result, it has been found that Marvin allows journalists to try various interactions with tools that can play an important role in their daily work. For this reason, Marvin is accepted as a successful approach to investigate the preferences of journalists about how they interact with robot journalism.

In their study, Caswell and Dörr (2018) introduced a computational approach to developing richer and more complex event-driven narratives rather than simple texts for automated journalism. The data models required to write event-driven narratives automatically are not sufficient. For this reason, a prototype database is defined for structured events and narratives, and then two methods are presented to use the event and narrative data in this database. Moreover, a framework was proposed to evaluate automatically generated event-driven narratives, technical and editorial challenges regarding the use of the approach in practice were discussed, and several high-level conclusions were opened for discussion on the importance of data structures in automated journalism workflows.

The distribution of 11 companies providing automatic content creation services by country is as follows: AX Semantics, Text-On, 2txt NLG, Retresco and Textomatic in Germany; Narrative Science and Automated Insights in the United States; Syllabus and Labsense in France; Arria is located in the UK, and Tencent is located in China. Yandex can be added to the list from Russia. Eight companies focus on providing content in one language, while the remaining four offer their services in multiple languages. On the other hand, these

companies do not see themselves as a journalistic institution. Therefore, their names are not associated with journalism, and their products are not intended to provide content in the field of journalism. On the contrary, their technology can be applied to any data from any industry (Graefe, 2016). Automated Insights, one of these companies, consists of robot journalists looking for interesting trends. For example, they examine data to focus on stories that only appeal to a small audience (such as the performance summary of a soccer team that a person is a fan of). Most journalists say that they want to write an article that will be read by many people, but the purpose of Automated Insight is to do the opposite: "We will create a million pieces of content that we hope only one person will read among one million people." (Rutkin, 2014, p. 22).

The system of the startup named Knowhere collects live news from hundreds of sites and produces three different versions by rewriting them: A positive, critical, and impartial perspective. According to the technology site Motherboard, this technology reviews thousands of news published on the internet every day, and news texts are checked by language recognition and word/sentence recognition algorithms. The screening is carried out in a way that the texts can distinguish between accurate information and biased discourses. Political-based news is categorized in a positive, critical, and impartial perspective. Knowhere artificial intelligence can produce three versions of the news about a 1 to 15-minute process after all analyses. Newly written versions are under the control of human editors and, if approved, are published in their news portals and mobile app called *Knowhere News*. Nathaniel Barling, co-founder and Chief Editor of Knowhere, states in the official statement about the system: "We are practicing a form of journalism that overcomes information overload and its resulting silos, attempting to reconcile the many different narratives spun out of every story, and taking our first steps towards a truly comprehensive and comprehensible source of record for all." (News Center, 2018)

Development of a Robot Journalism Application named Robottan Al Haberi

In this study, robot journalism application named Robottan Al Haberi developed produces news text by placing data on certain templates about weather, exchange rates, and earthquake. Robottan Al Haberi has been developed using the C# programming language in the Visual Studio 2019 integrated development environment. Data are provided by web service methods from electronic data distribution systems of the General Directorate of Meteorology, Central Bank, and Kandilli Earthquake Observatory respectively. The news texts, which are produced by placing the data collected externally from the institutions in the appropriate spaces on the template and with a maximum length of 280 characters, are automatically shared via the Twitter account @robottanalhaber. Weather information is shared once a day at 08:00; exchange rate information is shared three times a day at 08:00, 13:00, and 17:00. Earthquake information is updated every minute, and information about recent earthquakes is shared

instantly. The working model of the Robottan Al Haberi application consists of three stages (Figure 1). In the first stage, weather, exchange rates, and earthquake data are obtained from the electronic data distribution systems of the institutions, as mentioned earlier by the web service method and recorded on a database (data collection and storage stage). In the second stage, the quantity of the collected data is checked for different situations, and news texts are created in a way not to exceed 280 characters by placing data in pre-determined templates according to the size (data processing stage). In the third stage, the created news texts are shared on Twitter via the API platform at the hours and frequencies mentioned above (data sharing stage).

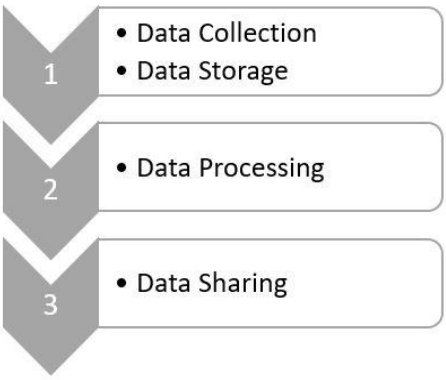

Figure 1. Working Model of Robottan Al Haberi Application

Data Collection and Storage

In the first stage of the application, three different functions were prepared, which can collect data from three different sources: General Directorate of Meteorology, Central Bank, and Kandilli Observatory. The GetWeatherForecastData function collects weather data at 08:00 once a day. The GetExchangeRatesData function collects exchange rate data three times a day at 08:00, 13:00, and 17:00. The GetEarthquakeData function works every minute to collect earthquake data. A timer object has been added to the program interface to control the running times of these functions.

A web service method was used as data collection method. The web service method is an electronic data-sharing model allowing data sharing between databases of different applications over the internet. In the web service method, it does not matter in which language or platform the systems that will exchange data are developed. These systems, which are developed and working independently from each other, can connect via hypertext transfer protocol (HTTP), which is the standard transfer protocol of the internet and can share data through extensible markup language (XML) outputs. As an example, the code fragment that collects data for the highest and lowest air temperature values in Istanbul is shown in Figure 2, and the code fragment that collects exchange rate data is shown in Figure 3.

```
string IstanbulMax;
string IstanbulMin;
string url = "https://www.mgm.gov.tr/FTPDATA/analiz/sonSOA.xml";
DataSet dsWeatherForecast = new DataSet();
dsWeatherForecast.ReadXml(@url);
DataRow dr = dsWeatherForecast.Tables[1].Rows[0];
DataRow dr2 = dsWeatherForecast.Tables[0].Rows[0];
IstanbulMax = dr[5].ToString();
IstanbulMin = dr[6].ToString();
```

Figure 2. Part of Sample Code Used to Collect Weather Data

```
dsExchangeRates = new DataSet();
filename = "http://www.tcmb.gov.tr/kurlar/today.xml";
dsExchangeRates.ReadXml(@filename);
DataRow dr = dsExchangeRates.Tables[1].Rows[0];
DollarSelling = Convert.ToDouble(dr[4].ToString().Replace('.', ','));
DollarBuying = Convert.ToDouble(dr[3].ToString().Replace('.', ','));
dr = dsExchangeRates.Tables[1].Rows[3];
EuroSelling = Convert.ToDouble(dr[4].ToString().Replace('.', ','));
EuroBuying = Convert.ToDouble(dr[3].ToString().Replace('.', ','));
```

Figure 3. Part of Sample Code Used to Collect Exchange Rate Data

The collected data is recorded in the application database. For example, the data set structure, which is used to save the earthquake data from the Kandilli Observatory in the application database, is shown in Figure 4.

					Magnitude			
Date	Time	Latitude	Longitude	Depth(km)	MD	ML	Mw	Location
2020.03.24	12:20:11	39.4337	27.8782	5.0	-.-	4.1	-.-	BEREKETLI-(BALIKESIR)
2020.03.24	11:59:05	38.2735	38.8080	11.2	-.-	1.8	-.-	KARSIYAKA-PUTURGE (MALATYA)
2020.03.24	11:54:52	37.1922	28.1635	0.0	-.-	1.8	-.-	CIFTLIKKOY-(MUGLA)
2020.03.24	11:39:32	37.2920	27.8520	22.5	-.-	1.8	-.-	TUZABAT-MILAS (MUGLA)
2020.03.24	11:24:06	39.1047	27.7975	9.8	-.-	3.2	-.-	KARAKURT-KIRKAGAC (MANISA)
2020.03.24	11:17:51	39.3872	25.8088	8.8	-.-	2.1	-.-	EGE DENIZI
2020.03.24	09:53:13	38.3495	38.7802	2.9	-.-	1.7	-.-	CANAKCI-KALE (MALATYA)
2020.03.24	08:20:31	37.9565	35.1385	1.6	-.-	2.8	-.-	KAVLAKTEPE-CAMARDI (NIGDE)
2020.03.24	08:05:57	38.3950	39.1820	12.8	-.-	1.6	-.-	KOSEBAYIR-SIVRICE (ELAZIG)
2020.03.24	07:03:09	38.3063	38.8398	3.4	-.-	2.0	-.-	GUNDEGER-PUTURGE (MALATYA)
2020.03.24	06:59:30	38.2985	38.8123	3.1	-.-	2.2	-.-	BOLUKKAYA-PUTURGE (MALATYA)
2020.03.24	06:58:28	38.6362	43.1563	4.1	-.-	2.7	-.-	MOLLAKASIM-(VAN)
2020.03.24	06:43:00	37.0332	28.5652	5.0	-.-	1.6	-.-	BALCILAR-KOYCEGIZ (MUGLA)
2020.03.24	06:33:24	37.4112	35.9048	16.6	-.-	1.2	-.-	COBANPINARI-KOZAN (ADANA)
2020.03.24	06:24:46	37.4670	35.9995	4.2	-.-	2.4	-.-	AKCATAS-SUMBAS (OSMANIYE)
2020.03.24	06:17:14	39.1220	27.8163	9.2	-.-	1.9	-.-	KARAKURT-KIRKAGAC (MANISA)
2020.03.24	06:13:42	38.4700	39.2227	3.3	-.-	0.8	-.-	ALINCIK-SIVRICE (ELAZIG)
2020.03.24	06:11:57	38.3875	39.1475	5.2	-.-	0.7	-.-	USLU-SIVRICE (ELAZIG)
2020.03.24	05:55:35	38.9498	26.9785	7.3	-.-	1.7	-.-	CANDARLI-DIKILI (IZMIR)
2020.03.24	05:51:16	39.0998	27.7860	8.9	-.-	3.0	3.0	ILYASLAR-KIRKAGAC (MANISA)
2020.03.24	05:28:01	38.4113	39.2657	10.0	-.-	1.6	-.-	DUZBAHCE-SIVRICE (ELAZIG)

Figure 4. Structure of the Earthquake Data Set Collected from Kandilli Observatory

Data Processing

In the second stage of the application, the obtained data is processed, and news texts are created. At this stage, the quantitative size of the information is checked for various situations, and news texts are created by placing the data in the pre-created templates according to the size values.

In Figure 5 and Figure 6, two sample code parts are given. In Figure 5, the piece of code that outputs three different texts for the news about the exchange rates is shown in three different times of the day. In Figure 6, there is a piece of code that outputs various news texts in cases where the earthquake intensity is less than 3, between 3 and 5, and above 5.

```
int time = DateTime.Now.Hour;
if (time < 13)
{
    textExchange.Text = "Merkez bankasından alınan bilgilere göre; Dolar fiyatları yeni güne " + DollarSelling + _
        & " TL seviyesinden, Euro fiyatları ise " + EuroSelling + " TL seviyesinden başladı.";
}
else if (time >= 13 && time < 17)
{
    textExchange.Text = "Merkez bankasından alınan bilgilere göre; Dolar kuru şu dakikalarda " + DollarSelling + _
        & " TL seviyesinde hareket ederken, Euro kuru ise " + EuroSelling + " TL seviyesinde bulunuyor.";
}
else
{
    textExchange.Text = "Merkez bankasından alınan bilgilere göre; Dolar kuru " + DollarSelling + _
        & " TL, Euro kuru ise " + EuroSelling + " TL seviyesinde günü kapattı.";
}
```

Figure 5. Part of Code That Writes Different Exchange Rate News in Three Different Time Periods

```
string earthquakesense = "";
if (Convert.ToDouble(dizi[5].ToString().Replace(".", ",")) < 3)
{
    earthquakesense = dizi[5].ToString() + " şiddetindeki deprem hissedilmedi.";
}
else if (Convert.ToDouble(dizi[5].ToString().Replace(".", ",")) >= 3 && Convert.ToDouble(dizi[5].ToString()) < 5)
{
    earthquakesense = dizi[5].ToString() + " şiddetindeki deprem panik yarattı.";
}
else if (Convert.ToDouble(dizi[5].ToString().Replace(".", ",")) >= 5)
{
    earthquakesense = dizi[5].ToString() + " şiddetindeki deprem korku yarattı.";
}
```

Figure 6. Code Part That Writes Different Earthquake News According to Earthquake Severity

In the application interface shown in Figure 7, there are examples of Turkish news texts resulting from the processing of the data. English translation of the texts is given in the Appendix.

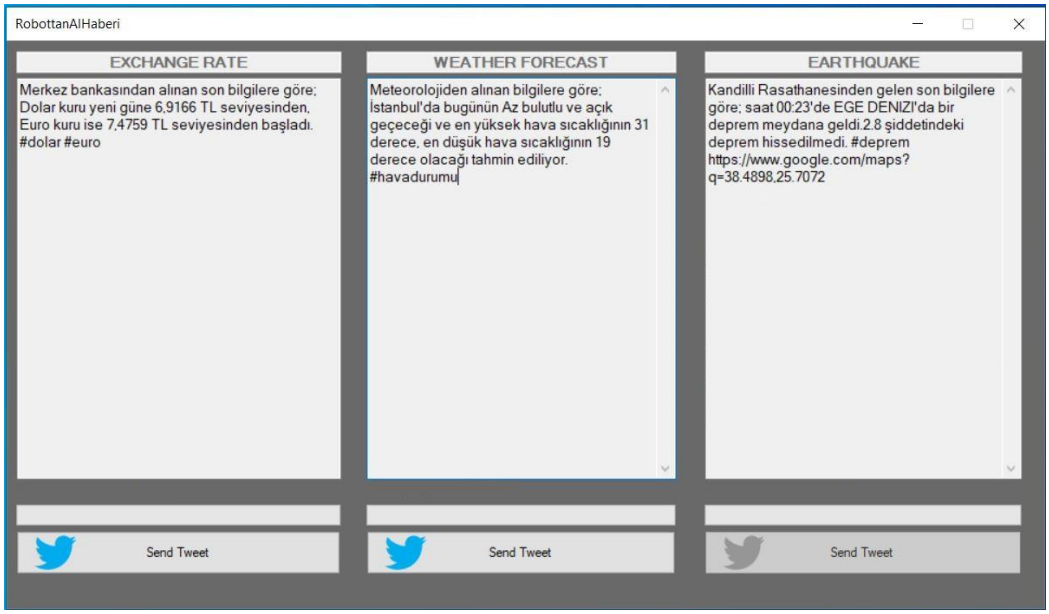

Figure 7. Application Interface

English translations of texts in Figure 7:

Exchange Rate: According to the data received from the Central Bank, the Dollar exchange rate started the new day at 6,9166 Turkish Lira (TL) and the Euro exchange rate at the level of 7,4759 TL. #dollar #euro

Weather Forecast: According to the data received from Meteorology; it is estimated that today, it will be low cloudy and clear in Istanbul and the highest air temperature will be 31 degrees and the lowest air temperature will be 19 degrees. #weatherforecast

Earthquake: According to the latest data from the Kandilli Observatory, an earthquake occurred in the Aegean Sea at 00:23. The magnitude of the earthquake was 2.8 and not felt. #earthquake

Data Sharing

As seen in Figure 7, the application interface consists of a single screen. The last tweets and sharing times for all three types of news are displayed in the area reserved for the title. Tweets are shared at certain times of the day through the timer running in the background. On the other hand, there are also the Send Tweet buttons that allow manual sharing in the application interface. If the timer shares the news displayed in the text field, the Send Tweet button is disabled and becomes inactive.

Sharing is occurred without a user (editor) involvement or approval by providing program-based access to the Twitter platform. Twitter APIs (application programming

interface) is used for this process. APIs are *talk* methods that allow computer programs to request and deliver information to each other (Twitter, 2020).

The requisite permissions must be obtained from Twitter to use these methods. For this study, application for authorization permission has been made to the Twitter Developer Platform on the internet address https://developer.twitter.com/, and the Twitter API Platform can be used for the @robottanalhaber account after obtaining the requisite permissions.

Conclusion

Robot journalism can be defined as algorithms that automatically convert them into news texts without human participation by using the data. This is a comprehensive statement. When the related literature is examined, it can be said that some approaches to implementation are only for a specific purpose. For example, some authors refer to robot journalism as news generation by placing structured data in specific news templates. In contrast, others consider it as a personal choice based on the analysis of the data and news generation. While some of them are based on revealing an important event and news generation on it, some of them introduced a computational approach to production for developing more complex event-driven narratives rather than simple texts. Finally, news can be generated from different perspectives by rewriting the current news collected from the sites to create a neutral journalism source.

In this study, news generation was performed by placing structured data on three different news templates. News templates were created by examining the words used in the related news and the sentence structures related to them. The robot journalism application Robottan Al Haberi generates news text about the weather, exchange rates, and earthquakes and automatically shares short news via its Twitter account named @robottanalhaber. However, the variety of topics can be increased. For example, the list of movies that will be released in the new week; places to be open and closed on public holidays; the result and winner of the digital lottery (with its location); the number of tourists entering the country (with which country they come from); daily summary of the Olympic results; videos with the highest number of views daily on Youtube, etc.

We predicted that the @robottanalhaber account would be paid attention and followed by the users since it delivers genuine and new information on three topics that interest news consumers. On the other hand, although Twitter is an interactive environment, such an account that offers only news to the user may not be expected to be interactive. In this case, the Twitter environment offers a single point to a multi-point communication model as in the case of traditional mass communication, therefore, our application can be evaluated as a newspaper's *breaking news* Twitter account.

The robot journalism approach in this study is a technology-reductionist approach. Future applications related to the subject will inevitably increase. The situation can be compared to a journalist gaining experience over time and producing better news texts using AI. In other words, robot applications will improve over time, and automated generation will evolve into a more comprehensive form. At this point, it can be suggested that human journalists should be more integrated with algorithms to convey similar experiences that they had in journalism.

References

Ali, W., & Hassoun, M. (2019). Artificial intelligence and automated journalism: Contemporary challenges and new opportunities. International Journal of Media, Journalism and Mass Communications, 5(1), 40-49. https://doi.org/10.20431/2454-9479.0501004.

Bucher, T. (2017). Machines don't have instincts: Articulating the computational in journalism. New Media & Society, 19(6), 918-933. https://doi.org/10.1177/1461444815624182.

Clerwall, C. (2014). Enter the robot journalist: Users' perceptions of automated content. Journalism Practice, 8(5), 519–531. https://doi.org/10.1080/17512786.2014.883116.

Dalen, V. A. (2012). The algorithms behind the headlines: How machine-written news redefines the core skills of human journalists. Journalism Practice, 6(5-6), 648-658.

Day, C. (2018). Robot science writers. Computing in Science & Engineering, 20(3), 101-101. https://doi.org/10.1109/MCSE.2018.03202638.

Dörr, K. N. (2016). Mapping the field of algorithmic journalism. Digital Journalism, 4(6), 700–722. https://doi.org/10.1080/21670811.2015.1096748.

Graefe, A. (2016). Guide to automated journalism. Columbia Journalism School. https://academiccommons.columbia.edu/doi/10.7916/D8QZ2P7C/download

Güz, N., & Yeğen, C. (2018). Bir dijital gazetecilik biçimi: Robot gazetecilik. Proceedings of the International Symposium on Communication in the Digital Age (pp. 328-339). Mersin, Turkey.

Jung, J., Song, H., Kim, Y., Im, H., & Oh, S. (2017). Intrusion of software robots into journalism: The public's and journalists' perceptions of news written by algorithms and human journalists. Computers in Human Behavior, 71, 291-298. https://doi.org/10.1016/j.chb.2017.02.022.

Kaa, H., and Krahmer, E. (2014). Journalist versus news consumer: The perceived credibility of machine written news. Proceedings of the Computation+Journalism Conference (pp. 1-4). New York.

Karaduman, M. (2017). Changing journalism and its new types. In E. Doğan & E. Geçğin (Eds.), Current debates in public relations, cultural & media studies (pp. 131-146). Ijopec.

Karlsen, J., & Stavelin, E. (2014). Computational journalism in norwegian newsrooms. Journalism Practice, 8(1), 34-48. https://doi.org/10.1080/17512786.2013.813190.

Kim, D., & Kim, S. (2017). Newspaper companies' determinants in adopting robot journalism. Technological Forecasting & Social Change, 117, 184-195. https://doi.org/10.1016/j.techfore.2016.12.002.

Latar, N. L. (2018). Robot journalism: Can human journalism survive?. World Scientific.

Lee, N., Kim, K., & Taeseon, Y. (2017). Implementation of robot journalism by programming custombot using tokenization and custom tagging. Proceedings of the 19th IEEE International Conference on Advanced Communications Technology, 566-570. Pyeongchang, Korea.

Lindén, C. G. (2017). Algorithms for journalism: The future of news work. The Journal of Media Innovations, 4(1), 60-76. https://doi.org/10.5617/jmi.v4i1.2420.

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

Mayes, R. (2014). The future of futurists: Can a machine produce this forecast? The Futurist, 48(6), 21-23.

McCartney, P. (2015). Robotic journalism and nursing. MCN: The American Journal of Maternal/Child Nursing, 40(5), 330. https://doi.org/10.1097/NMC.0000000000000175.

Melin, M., BäCk, A., SöDergåRd, C., Munezero, M. D., LeppäNen, L. J., & Toivonen, H. (2018). No landslide for the human journalist - An empirical study of computer-generated election news in finland. IEEE Access, 6, 43356-43367. https://doi.org/10.1109/ACCESS.2018.2861987.

Montal, T., & Reich, Z. (2017). I, robot. You, journalist. Who is the author? Digital Journalism, 5(7), 829-849. https://doi.org/10.1080/21670811.2016.1209083.

Monti, M. (2018). Automated journalism and freedom of information: Ethical and juridical problems related to AI in the press field. Opinio Juris in Comparatione, 1(1), 1-20.

Narin, B. (2017). Uzman görüşleri bağlamında haber üretiminde otomatikleşme: Robot gazetecilik. Galatasaray Üniversitesi İleti-ş-im Dergisi, 27, 79-108. https://doi.org/10.16878/gsuilet.373242.

Ombelet, P. J., Kuczerawy, A., & Valcke, P. (2016). Employing robot journalists: Legal implications, considerations and recommendations. Proceedings of the 25th International Conference Companion on World Wide Web (pp. 731-736). Montréal Québec.

Rutkin, A. (2014). Machines write the news. New Scientist, 2962, 22.

Sarılar, N. B. E. (2019). Robot journalist or human journalist?: An analysis is over news articles. Proceedings of the Communciation and Technology Congress (pp. 201-208). Istanbul Turkey.

Shekhar, S. (2016). Robot journalism: The advent of high tech storytelling. PCQuest, April, 20-22.

Sim, D. H., & Shin, S. J. (2016). Implementation of algorithm to write articles by stock robot. International Journal of Advanced Smart Convergence, 5(4), 40-47.

Szews, P. (2018). Data journalism, geojournalism, CAR i robot journalism jako nowe odmiany i terminy w dziennikarstwie. Acta Universitatis Lodziensis. Folia Litteraria Polonica, 51(5), 209-222.

Tatalovic, M. (2018). AI writing bots are about to revolutionise science. Journal of Science Communication, 17(1), 1-7. https://doi.org/10.22323/2.17010501.

Túñez-López, J. M., Toural-Bran, C., & Cacheiro-Requeijo, S. (2018). Uso de bots Y algoritmos para automatizar la redacción de noticias: Percepcıón Y actıtudes de los periodistas en España. El profesional de la información, 27(4), 750-758. https://doi.org/10.3145/epi.2018.jul.04.

Twitter. (2020). About Twitter's APIs. Help Center. https://help.twitter.com/en/rules-and-policies/twitter-api

Vállez, M., & Codina, L. (2018). Computational journalism: Evolution, cases and tools. El profesional de la información, 27(4), 759-768. https://doi.org/10.3145/epi.2018.jul.05.

Verhulst, J., Rofman, M., Heussen, N., Geer, C., & Vos, K. A. J. (2017). MARVIN: An interactive robot journalism simulation. https://pdfs.semanticscholar.org/702a/240e1dd1d749b580c39e59c523da8b4d7af2.pdf?_ga=2.31931390.1598439443.1589498673-637886212.1589498673

Waddell, T. F. (2019). Attribution practices for the man-machine marriage: How perceived human intervention, automation metaphors, and byline location affect the perceived bias and credibility of purportedly automated content. Journalism Practice, 13(10), 1255-1272. https://doi.org/10.1080/17512786.2019.1585197.

Wölker, A., & Powell, T. (2018). Algorithms in the newsroom? News readers' perceived credibility and selection of automated journalism. Journalism, February, 1-18. https://doi.org/10.1177/1464884918757072.

Young, A. (2016). All Things Digital in 2016. E&P Editor & Publisher, 149(2), 50-56.

Zheng, Y., Zhong, B., & Yang, F. (2018). When algorithms meet journalism: The user perception to automated news in a cross-cultural context. Computers in Human Behavior, 86, 266-275. https://doi.org/10.1016/j.chb.2018.04.046.

Tornoe, R. (2014, September 23). Learn to Stop Worrying and Love Robot Journalists. Editor&Publisher. https://www.editorandpublisher.com/columns/digital-publishing-learn-to-stop-worrying-and-love-robot-journalists/

Birer, G. C. (2016, May 25). Bir bilgisayarın roman yazdığı gün. Tübitak Bilim Genç. https://bilimgenc.tubitak.gov.tr/makale/bir-bilgisayarin-roman-yazdigi-gun

İrvan, S. (2017, August 7). Robot gazeteciler geliyor. Yeni Medya ve Gazetecilik. https://suleymanirvan.blogspot.com/2017/08/robot-gazeteciler-geliyor.html

Loosen, W. (2018, March 18). Four forms of datafied journalism. Journalism's response to the datafication of society. Communicative Figuration Working Paper. https://www.kofi.uni-bremen.de/fileadmin/user_upload/Arbeitspapiere/CoFi_EWP_No-18_Loosen.pdf

News Center. (2018, April 20). Knowhere yapay zekası etik haberciliği geri getirecek. xTRlarge. https://www.xtrlarge.com/2018/04/20/knowhere-yapay-zeka-etik-habercilik

Caswell, D., & Dörr, K. (2018). Automated journalism 2.0: Event-driven narratives. Journalism Practice, 12(4), 477-496. https://doi.org/10.1080/17512786.2017.1320773

Bibliographic information of this paper for citing:

Tosyalı, Hikmet; Aytekin, Çiğdem (2020). Development of Robot Journalism Application: Tweets of News Content in the Turkish Language Shared by a Bot. *Journal of Information Technology Management*, Special Issue, 68-88.