언론정보학석사 학위논문

The Frames and the Emotions in the News Media’s Coverage about Artificial Intelligence

인공지능 관련 뉴스 기사의 프레임, 감정 분석

2022년 8월

서울대학교 대학원

언론정보학과

김 유 정

The Frames and the Emotions in the News Media's Coverage about Artificial Intelligence

지도 교수 이 철 주

이 논문을 언론정보학석사 학위논문으로 제출함

2022년 5월

서울대학교 대학원

언론정보학과

김 유 정

김유정의 언론정보학석사 학위논문을 인준함

2022년 7월

위 원 장 ______ 한규섭 ______ (인)

부위원장 ______ 황현정 ______ (인)

위 원 ______ 이철주 ______ (인)

Abstract

The Frames and the Emotions in the News Media's Coverage about Artificial Intelligence

김 유 정(YuJeong Kim)
언론정보학과(Department of Communication)
The Graduate School
Seoul National University

This study examines how artificial intelligence (AI) is presented in the news media by examining the frames and emotions expressed in news coverage about AI. For analysis, I used computational text analysis techniques –structural topic model (STM) to extract frames and NRC Emotion Lexicon and Linguistic Inquiry and Word Count (LIWC) to detect emotions. Then I examined their correlations with the political ideology of media outlets (conservative vs. liberal) and media type (newspapers vs TV news). By identifying the frames and the emotions embedded in the news media, it would be possible to predict how they influence the formation of public opinions and attitudes towards AI.

Keywords : artificial intelligence, news media, topic modeling, computational analysis, frame, discrete emotion
Student Number : 2019-25994

Table of Contents

Chapter 1. Introduction

1. Study Background

Today, artificial intelligence (AI) has become one of the key factors of socioeconomic transformation. The advancement in AI technology is expected to impact nearly every sector of the society, including transportation, education, labor market and medical industry. Due to its large spectrum of influence, however, AI application has raised major discourses related to various fields such as human values, privacy, safety and accountability. (OECD, 2019)

Despite the multitude of social discussions and questions AI has given rise to, AI technology itself still remains in the early stage of development – that is, AI has just begun to be applied to people’s everyday lives. Accordingly, the lay public remains largely unfamiliar with this new technology. While the term ‘Artificial Intelligence (AI)’ itself does frequently appear in daily discourses, the public still lacks clear understanding of what AI exactly is. For instance, the concept of AI is commonly confused with and mistaken for robots. In a similar vein, AI has not fully entered the realm of public discourse. Zhang and Dafoe (2019) point out that when compared with the relatively active discussions that happen in technological or political sectors about AI and its application, the public has not been able to participate in the process of shaping these conversations.

This indicates that public opinion on artificial intelligence is in much flux, as it is still in the early stages of the issue cycle. That is, at this point of the issue cycle, public perceptions and attitudes towards AI are variable, susceptible to influences from various factors within cognitive and affective processes (Lee et al., 2005).

This study investigates how the legacy news media uses the frames and the emotions in delivering AI-related news, focusing on the differences depending on the news agency's political viewpoints (conservative vs. liberal) and the type of media (TV vs. newspaper). It collects US data and analyzes them using two computational methods – NRC Emotion Lexicon and LIWC.

Chapter 2. Literature Review and Research Aim

1. Public opinion towards emerging technologies

Public opinion has been deemed as an important field of research in science communication because they shape policy decisions. There is a vast amount of literature that investigate public opinion on emerging science technologies. Previous science communication research has focused on public attitudes and perceptions towards controversial science issues such as climate change, nanotechnology, or genetically modified food. Caughey and Warshaw (2018) show how public opinion in the United States played its role in forming policy outcomes, and how policies were shaped in response to the public's political preferences. Page and Shapiro (1992) identified the reason behind such tendency and stated that policy makers are prone to be more likely to respond to policy directions that are more favored by the mass. That is, policy makers may not entirely depend on public opinion when implementing certain policies, but they may be more likely to take actions when the public is in favor of such policy decisions (Page & Shapiro, 1992). Sometimes, it is also possible for the public to have direct influence on policies (Cobb, 2005). For example, the voters in California approved a plan to fund a decade-long stem cell research in 2004 (Holden, 2004). The public and its opinion, therefore, do matter in policy-making decisions. As observed in policy domains such as free trade, national security or environmental issues, it is reasonable to expect the public to exert more influence over time in shaping policies related to artificial intelligence (Zhao &Dafoe, 2020). Hence, it is imperative to pay close attention to how the public perceives and feels about AI and which factors influence such attitudes.

Previous studies have highlighted the role of the mass media as the contributor in the formation of public perceptions and attitudes toward science. Nisbet et al. (2003) found that the news media and its coverage on scientific controversies, in particular, have played a vital and interactive role in such context. Nisbet and his colleagues (2013) argue that the mass media forms a public arena where scientific issues are presented to various individuals, interest groups, and decision makers within a society. Moreover, the media plays a crucial role in shaping how scientific issues or controversies are defined, discussed, and symbolized (Anderson, 2014).

In explaining the formation of attitudes and perceptions towards emerging technologies, the traditional scientific literacy models have mainly focused on information deficits of the lay public (Miller, 1998). In comparison, the 'cognitive miser model' proposed by other researchers (Fiske & Taylor, 1991; Popkin, 1994) suggests that people collect little or no information about a given issue and rather tend to make decisions based on limited amount of information. Considering the limited time and energy ordinary people can spend on obtaining in-depth understanding of scientific discoveries, it is natural and rational for them to rely on various shortcuts and sources to aid their judgement (Popkin, 1994). In this process, the media and its portrayal of emerging technologies serve as one of the most important sources people depend on (Scheufele & Lewenstein, 2005). In case of emerging technologies, a vast majority of the public usually holds only limited direct experiences and knowledge, making news coverage even more suitable to function as a major heuristic. As previously mentioned, artificial intelligence is regarded as an emerging science technology that the public is still unfamiliar with. It is therefore reasonable to expect ordinary citizens to depend on the media and news to form judgements on AI and relevant issues surrounding it. Hence, it is necessary to examine the media portrayal of artificial intelligence in order to check how the news media outlets deliver AI-related information and ultimately participate in opinion formation.

Traditionally, previous literature on media effects has mainly focused on the cognitive effects the media has on the public's attitudes (Potter & Riddle, 2007). However, recent researches also shed light on the emotional influences of the media, acknowledging the fact that affective aspects interfere with how people perceive science or emerging technologies. In other words, while cognition –

such as scientific knowledge and facts – contributes towards opinion formation about scientific issues, it alone cannot fully explain the dynamics that exist in how individuals perceive and form attitudes about such issues. For example, Yang and Chu (2018) conduct research on how discrete emotions such as fear, anger and sadness are related to the US public's perception of the Ebola outbreak. Kuhne and Schemer (2015) examine how anger and sadness influence individual opinion formations and information processing. In short, cognition and emotion work together to affect one's perception or information processing.

Some researchers even focus on how cognitive and affective aspects jointly interact with each other to exert influence on information processing and decision making. For example, Lee et al. (2005) analyze survey data on people's attitudes about nanotechnology, postulating that cognitive and affective factors simultaneously work in tandem to shape how people form their opinions on emerging technologies. Acknowledging the roles cognition and emotion play in forming public opinion, this study covers both cognitive and affective aspects of the media influence by identifying cognitive frames and emotions embedded in news stories about artificial intelligence.

2. Public opinion towards artificial intelligence

Artificial intelligence (AI) is defined as "the development of machines capable of sophisticated (intelligent) information processing" (Dafoe, 2018, pg5). AI is capable of independently – that is, without human instructions – making decisions or performing actions that commonly demand human intelligence (Zhang & Dafoe, 2019).

There are several researches that examine public perception towards AI and its governance. Zhang and Daofe (2019) conduct a nationally representative survey with 2,000 American adult participants, asking their opinions on issues such as their level of trust in the parties developing and regulating AI, workplace automation, AI governance and its potential impacts and international cooperation in AI development and governance. The results show that there are more Americans who support than oppose the development of artificial intelligence. There are great demographic differences in the responses – wealthy, educated males with higher level of experience with technology are more prone to be in support of developing AI. Meanwhile, results show that there are also public concerns towards AI. More than 80 percent of the participants agreed that AI should be managed carefully and answered that AI will eliminate jobs rather than create them. They also expressed doubt that high-level machine intelligence will have positive impact on humanity in the future. Wilson and his colleagues (2020) conduct a national survey in the United States. The survey results reveal that the American public in general holds favorable attitudes toward AI, expecting its positive influence on various sectors such as the job market or health care. At the same time, there were also wary voices with concerns over job losses, privacy or cyber security issues.

While some researches seek to gain an understanding of public attitude towards AI through surveys, others analyze discourses about artificial intelligence expressed in the mass media. By examining articles about AI in the New York Times over the past thirty years, Fast and Horvitz (2017) find that the public discussion about AI has become more prominent than the past, especially after 2009, following the start of a wide use of deep learning. They also maintain that the discussions found in the New York Times have shown consistent tendency of being more optimistic than pessimistic. Garvey and Maskal (2020) examine the New York Times and articles collected by a news aggregator 'News API' to conduct a sentiment analysis on news stories about artificial intelligence. This research reveals that, contrary to the popular belief that news media outlets portray AI in a negative view, they actually tend to be more positive than negative.

Although it is true that there are previous researches on public perception about AI, the topic is still relatively unexplored. As AI itself is a new technology that only just recently entered the everyday lives of the lay public, there are only a limited number of related literature at the moment. A big majority of researches conduct surveys (Merenkov et al., 2020; Wilson et al., 2020; Zhang & Dafoe, 2019), directly asking people their attitudes and perceptions towards AI. Only few researches investigate the mass media coverage on artificial intelligence and its impact on public opinion, and they only focus on cognitive and emotional sectors separately.

In sum, none of the past researches simultaneously observe the frames and the emotions displayed in news coverage about artificial intelligence. As mentioned earlier, the news media is an important channel for public engagement in science and emerging technology, as it is capable of influencing people's perception and attitudes toward different technologies and their applications. As also discussed above, cognition and emotion work together when shaping opinions and attitudes towards particular issues. Hence, by examining the frames and emotions that are used to describe artificial intelligence in the news media, this study obtains a more comprehensive understanding of various discourses surrounding AI technology and the public opinion towards it and thus gain valuable insights.

Chapter 3. Conceptual Framework

1.1. Frames

The way the media covers an event or an issue – commonly regarded as "framing" – has received considerable spotlight in the academia due to the understanding that frames are capable of shaping public opinion. In fact, framing has been regarded as one of the most "central, applicable and contested theories" in communication research (Walter& Ophir, 2019).

The concept of framing has been widely discussed in the field of communication, across diverse disciplines such as sociology or psychology. Gamson and Modigliani (1981) state that a frame is a "central organizing idea or storyline that provides meaning to an unfolding strip of events, weaving a connection among them" (p.375). There are different types of framing effects that actually pertain to different phenomenon but share identical labeling (Cobb, 2005). In this study, framing refers to 'issue framing', which is more commonly employed in political communication studies as opposed to 'equivalency framing'. Issue framing means that when describing an identical issue, "qualitatively different yet potentially relevant considerations" are used (Druckman, 2004, p672). Entman (1993) suggests that issue framing refers to the act of "selecting some aspects of a perceived reality and making them more salient in a communicating context in such a way as to promote a particular problem definition, casual interpretation, moral evaluation, and/or treatment recommendation" (p.52). That is, a frame infuses a perspective into a message and raises the salience of a particular information, and this process may influence receivers' views (Nabi, 2003). In short, framing is about how a particular information is being presented in public discourse, rather than about what is being presented (Scheufele & Iyengar, 2014). This definition indicates that the way issues or events are portrayed and discussed in news stories may affect how receivers understand them.

Framing effect can be understood in regard to the heuristic processing of information. As mentioned earlier, individuals cannot have a holistic understanding of the world and must inevitably interpret issues and events happening around them in order to make sense of the world. As Goffman (1974) posits, individuals therefore employ interpretive schemas to more efficiently process new information and interpret them. That is, when it comes to understanding complex and unfamiliar issues such as information on science or technology, frames help them resonate with the existing cognitive schemas of the audience so that they can more easily and efficiently process the information (Scheuefele & Tewksbury, 2007). In other words, while the lay public may lack concrete factual information about new technologies, they will still use cognitive processes, such as frames, to form their own perceptions and judgments (Lee et al., 2015).

There is extensive literature on news frames and their effects on opinion formation of science and emerging technologies. For instance, Gaskell and his colleagues (1999) maintains that media framing of genetically modified foods (GMF) was associated with the increase of negative opinions amongst Europeans. Scheufele and Lewenstein (2005) postulate that cognitive shortcuts such as media frames affect how the public perceives and understands nanotechnology and can even influence the level of support for funding. Such researches underline the need to identify and examine what kinds of frames are used in news stories to depict the subject of interest. Acknowledging such necessity, the first research question is put forth as follows:

RQ1: What frames are salient in American news coverage of AI?

1.2. Frame building surrounding artificial intelligence

Frame building is a term coined to explain the way the news media select certain frames when presenting issues and the different social and structural factors or journalistic practices that contribute towards influencing frames (Kim et al., 2010). The process of building frames can be affected by various internal and external aspects of news agencies – for example, sociocultural norms and values, features of individual journalists, organizational orientations and pressures, and intervention of interests groups. (Scheufele, 1999).

One of the main factors that influence frame building process is the political orientation of news organizations. The political ideologies that each news organization pertain to are often reflected in the overall tone of the news stories and editorials. For example, conservative media naturally tend to emphasize individual freedom and limited regulation of the market while liberal media focus on the role of governmental intervention and regulation.

With artificial intelligence being in the stage of development, there are still much controversies over its governance and the legislation of related policies. Some of the main areas of contention include privacy, security or ethical issues. For example, there are voices of concern that AI might pose serious threat to privacy. The collection of large amount of data required for machine learning, the use of cloud computing, or the process of extracting knowledge from big data set allegedly bear possibilities of breaching privacy (Li & Zhang, 2017). Security is yet another issue that is commonly addressed – there are concerns that technology abuse, technical defects, or the possible self-awareness of intelligence may cause security problems (Li & Zhang, 2017). In addition, the data acquired for the development of AI is claimed to reflect biases of the real world, hence is said to hold danger of being discriminatory towards minorities. Due to such issues surrounding AI technology, AI and its governance have been in the center of debate for policymakers. As previously mentioned, conservatism tends to encourage free market and limited government regulations. Accordingly, conservative media may be more likely to emphasize the advantages and positive prospects that could be brought by the development of artificial intelligence – in this case, mostly the economic benefits AI technology can promise. On the other hand, in contrast to conservative media, liberal media may be more likely to call for regulations or measures to prevent the possible problems that could be caused by the application of AI. Based on these arguments, this study proposes the following hypotheses:

> H1a: Conservative media are more likely than liberal media to involve frames discussing economic benefits that can be brought by artificial intelligence.

> H1b: Liberal media are more likely than conservative media to involve frames discussing measures to prevent problems that can be caused by artificial intelligence.

This study also investigates if the type of media outlet (e.g., newspapers vs. TV) has an impact on the way news media discuss artificial intelligence. The professional practices of journalists in the process of news production are likely to affect how news stories are made. For instance, the use of episodic framing is one of the frequently-used professional practices in journalism. The use of episodic framing is more prominent in news stories on television than those in newspapers, mainly due to the audiovisual features of TV

(Iyengar, 1991). Meanwhile, newspaper articles are generally believed to deliver more amount of information in greater detail and provide more background context than television news (Eveland Jr et al., 2002). These differences stem from the fact that, as compared to newspapers, TV news has more limitations in terms of the time and space that could be used to deliver information – the time allocated for news broadcast tends to consist fewer words or stories than the length of newspapers does (Druckman, 2005). In relation to such differences caused by the medium of delivery, this study investigates whether frames found in news coverage about AI are contingent upon the type of media. Hence, the following research question can be proposed:

> RQ2: Are there meaningful differences in the frequently used frames in news coverage about AI between newspapers and TV news?

Considering the gap in the amount and depth of information presented in news articles between newspapers and television news, it is possible that economic issues may be more prominent in newspapers than in TV news. News coverage on economy or business often require more detailed information and are commonly accompanied by explanations about background context surrounding the issue. Based on the understanding of such journalistic practices, this study proposes the following hypothesis:

> H2: Newspapers are more likely than TV news to express frames discussing economic discourses about AI.

2.1. Emotion

Emotions are generally defined as "internal mental states representing evaluative reactions to events, agents, or objects that vary in intensity" that are "short-lived, intense and directed at some external stimuli" (Nabi, 2002, p. 289-290).

There are two basic models that are generally used when approaching the concept of emotion in the field of communication related studies – dimensional and discrete. The dimensional approach sees emotion as a motivational state that is shaped by arousal and valence consisting affective dimensions (Nabi, 2010). In other words, this approach characterizes emotion into two broad dimensions – the level of activation (arousal) and the state of pleasure/displeasure (valence). In contrast, the discrete emotion approach mainly focuses on categorical state of emotions which are characterized with unique cognitive appraisals, experiences and particular action tendencies (Izard, 1997). This approach holds that particular thought patterns about the environment surrounding one's goals lead to particular behavioral tendencies (Nabi, 2010). The two perspectives share a common grounding in that they both assess valence and intensity. However, as the discrete emotion approach takes into account the unique emotion states, it allows for more accurate predictions and explanations of human action and is thus more suitable for studying communication phenomenon (Nabi, 2010). The cognitive functional model (CFM) by Nabi (1999), which is one of the theoretical backbones of this study, is also based on the discrete emotion approach, assuming that emotions are linked with different goals and actions aimed at achieving such goals. This study therefore employed the discrete emotion approach when conceptualizing the concept of emotion.

Researches on the framing effects of emotions mostly build on cognitive appraisal theories (De los Santos & Nabi, 2019), which deem emotions as formed by how individuals appraise environmental stimuli (Lazarus, 1991) in association with different cognitive dimensions such as controllability, responsibility, pleasantness or relevance (Roseman, 1984; Smith & Ellsworth, 1985). In communication studies, the cognitive appraisal theory implies that news readers will appraise issues or events in a similar way as depicted in the news (Kuhne & Schemer, 2015).

As mentioned earlier, emotions account for an important part in the media effect studies because they can affect information processing and opinion formation. Nabi (1999) proposes the cognitive functional model (CFM) to explain how emotions induced by messages influence information processing and persuasive effects. The CFM (Nabi, 2002) holds that "a message evokes an emotion if its content reflects the emotion's core relational theme and if the receiver recognizes both that theme and its personal relevance" (p. 205). Such response leads to motivations to attend to or avoid the stimulus that induce the emotion, or to satisfy the goal induced by the emotion. Along with the initial response to the stimulus that evoke emotions, receivers are aroused with motivations to act consistently with the induced emotion to deal with problems encountered in the situation. In the field of communication, this implies that when a particular emotion is elicited by a news story, the processing, judgement, or perception of the story can be biased by the initial emotion (Nabi, 1999, 2002). Such understanding of message-relevant emotions is critical as it indicates that news media can also induce emotions that influence information processing, attention and accessibility, which ultimately lead to the formation of perceptions or judgements that are correspond to the emotion (Nabi, 2007).

While the vast majority of previous research focuses on cognitive elements of news framing effects, there are several studies that examine the emotionally-evocative frames spotted in news articles (De los Santos & Nabi, 2019). These studies mostly investigate emotions elicited by different news frames.

Emotions have become an important topic in the field of science communication in that they are perceived to take part in predicting individuals' attitudes about various social issues. Following such perspective, numerous studies have examined the relationship between emotions and public opinion or attitude. For example, Nabi (2003) explores the framing effects of emotion, maintaining that discrete emotions differentially affect information accessibility and information seeking and even support for policies on specific issues. Smith and Leiserowitz (2014) contend that the discrete emotions evoked when made to think about global warming were stronger predictors of related policies when compared with other sociodemographic variables. Nabi et al. (2018) examine the role of fear and hope in influencing attitudes and advocacy of climate change policies. The study by Kim and Cameron (2011) investigates emotional news frames in relation to corporate crisis. It reveals that emotional news frames differently affect people's information processing and how they perceive the company. Such literatures support the theoretical proposition that message-relevant emotions induced by news frames can form and affect opinions and judgments.

Despite the findings that could be brought by examining discrete emotions in the news media, the discrete emotions about AI is an under-explored field of research. There are previous studies that have taken a valence-based approach (positive vs. negative) in examining news coverage about AI (e.g. Garvey & Maskal, 2020), but none that explore discrete emotions. This study, therefore, raises the following research question in order to investigate which emotions have been used to describe AI in the news media. In regards to the differences in frames in the news media that are expected to stem from the differing political ideologies between liberalists and conservatives(RQ2), this study examines whether the same tendency can be spotted in the emotions embedded in news stories. Hence, this study proposes the following research question:

> RQ3: Are there meaningful differences in the discrete emotions expressed in news media about AI between conservative media and liberal media?

Though there are various discrete emotions that can be used, this study focuses on three - fear, anger and hope. This is due to the fact that they are percieved as a spectrum of emotions that hold different motivation goals, and more importantly, that all three of these emotions are most commonly emphasized in news coverage (De los Santos & Nabi, 2019). Below is an explanation of the three emotions discussing distinct conditions that arouse them and what influences they have on attitude formation.

Anger

Anger refers to a negative type of emotion that is caused by the perception of offenses (Lazarus, 1999) that work against me and mine. That is, anger is often induced when one encounters obstacles that are perceived to be in conflict with behaviors that are goal-oriented or threaten the interests of oneself or loved ones (Lazarus, 1999). This emotion often leads to desires to punish the perpetrator that is perceived to deliver such demeaning offenses (Nabi, 2003). In terms of the news media, news events intrinsically attribute blame or responsibility - for instance in the case of governmental scandals or harm-inducing negligence - which are likely to elicit anger (De los Santos & Nabi, 2019).

Fear

Fear is an emotion that is caused by the perception of close physical or psychological danger (Lazarus, 1991; Nabi, 1999). Such threats include variety of factors ranging from biological factors to sociocultural or individual influences. Stemming from the desire to seek protection, fear leads individuals to behave so that one can escape from the threatening factors to avoid the harmful situation (De los Santos & Nabi, 2019; Nabi, 1999). Journalists' selection of news tends to gear towards deviant or threatening events and ideas (Shoemaker, 1996) such as accidents or crimes, and this may result in arousing fear amongst its readers (De los Santos & Nabi, 2019).

Hope

Hope is elicited when one yearns for a wanted outcome in an uncertain situation (Lazarus, 1999). In other words, hope is induced in situations where one feels it is possible to gain what they desire, but when that success is uncertain (Smith & Ellsworth, 1985). It is based on a future-oriented feeling that includes a positive visualization of future outcomes. For instance, news articles that cover stories about people overcoming difficulties or major breakthroughs to certain problems may elicit hope (De los Santos & Nabi, 2019). Hope is likely to result in behaviors that drives one toward his/her desires and is linked with perseverance when faced with hardship (De los Santos & Nabi, 2019)

In line with the aforementioned ideological differences that liberalists and conservatives most likely have, it is possible to expect that similar tendencies may be reflected in the emotions that are used to describe AI. That is, conservative media which are more likely to shed light on the possible economic benefits brought by the development of AI may also be more likely to use hope to depict AI, for the purpose of forming positive images in the readers' minds. In contrast, the liberal media which are more likely to call for policy interventions to regulate AI technology may be more likely to use fear and anger, so as to arouse alarm against the new technology. Based on such understanding, this study proposes the following hypotheses:

> H3a: Conservative media are more likely than liberal media to express hope in AI related articles.

> H3b: Liberal media are more likely than conservative media to express fear and anger in AI related articles.

The media has often been blamed for forming and amplifying public concerns about emerging technologies. The news coverage about AI has not been an exception. For example, Garvey and Maskal (2020) claim that the news media has been allegedly using negative imagery from the movie *Terminator* to depict artificial intelligence, thereby engaging in arousing concerns amongst the lay public against AI. Meanwhile, although it is true that newspaper and TV may deliver identical news contents, there are differences in the style of delivery that stem from the differences in the type of media. This study, therefore, raises the next research question in order to investigate whether the variation by medium of delivery distinguishes emotions expressed in news.

> RQ4: Are there meaningful differences in the discrete emotions expressed in news media about AI between newspapers and TV news?

In comparison to print media, televised media tend to have higher emotional influences (Wanta, 1997). As a result, TV news are more likely to select and construct stories in a way that emphasize the necessary emotion in a relatively short length of time allocated per story (Driedger, 2007). In addition, due to the audiovisual features that are exclusive to televised media, TV news are also more prone to be dramatic, emotionally stimulating or sensational, as the popular credo of local TV news stations "If it bleeds, it leads" suggests (Cooper & Roter, 2000). For example, Cho and his colleagues (2003) analyze news coverage on September 11th terrorist attack in the US to discover that television news was consistently more emotional than printed news. The same line of idea can be applied in the case of news stories about AI. The *Terminator* imagery or the machines-take-over-humans narrative are more intuitive and dramatic, as opposed to the economic or technological benefits that could be brought by the development of artificial intelligence. In other words, stories that appeal to fear and anger may be easier to be dramatized or more suitable to leave needed emotions in the two to three minutes given to each story on TV. Based on the theoretical understandings and evidence above, I propose the following hypotheses.

H4a: TV news are more likely than newspapers to express fear and anger when discussing AI.

H4b: Newspapers articles are more likely than TV news to express hope when discussing AI.

Chapter 4. Methods

This study discovered frames and emotions infused in the news media's articles by examining both printed and TV news in the United States from 2019 to 2021 (from January 1st, 2019 to December 31st, 2021). In order to embrace a broad and balanced spectrum of American news media, this study collected articles from various media outlets. As for newspapers, the top four most circulated daily newspapers were collected –The Wall Street Journal, USA today, The New York Times, Los Angeles Times. As for TV news, this study followed Ksiazek et al. (2019) selection of TV news programs which include two main cable news networks (Fox News and CNN) and the three major broadcast networks (ABC, NBC, CBS). In regard to the categorization of each news media’ s political ideologies, this study referred to previous studies (Budak et al., 2016; Flaxman et al., 2016). Wall Street Journal and Fox News was categorized as conservative media, the New York Times and CNN as liberal, and the rest as neutral.

1. Frame analysis

The analysis of frames includes two steps – first, the identification of frames used in the news media when covering certain issue and second, the coding of the frames that are present in articles (Burscher et al, 2016). The first step includes the process of detecting and defining frames in news, and the second step entails the annotation of frames identified in the first step. In order to find frames in the news media articles, this study used STM topic modeling employing a statistical computing program ‘R’.

In many of the previous researches on framing, news frames have been coded with manual content analysis where human coders are trained to manually code frame indicator questions (Burscher et al., 2016). The clear downside of human coding is that it is time-and money- consuming and has high possibility of being biased. In this sense, the advent of computer technology gives rise to new methods to tackle the limitations of human coders. Computers are capable of processing large amounts of data and repetitively coding frames with higher speed and lower costs. Due to such advantages, many researches have employed computers in analyzing news frames.

1.1 Topic Modeling – Structural Topic Model

One of the most popular computational methods for identifying frames is topic modeling. Topic modeling refers to an unsupervised machine learning method that uses a Bayesian approach to analyze text. It is used to quickly and efficiently detect hidden thematic structure in large amount of text documents (DiMaggio et al., 2013; Maier et al., 2018). In comparison to dictionary–based approach – a previously widely–used computer aided technique for coding news frames that required researchers to manually design pretest and refine search queries to build models for classifying contents – topic model overcomes the biases caused by the subjective conceptions and limited knowledge of the researcher in the process of manual intervention (Burscher et al., 2014).

Latent Dirichlet Allocation (LDA) is one of the simplest (Blei, 2012) and most widely used (Guo et al., 2016) techniques of topic models. It is capable of identifying latent topics based on the distribution of words presented in documents rather than predefined categories (Kwon et al., 2019). This characteristic contributes to its higher validity when compared with previous techniques (Guo et al., 2016). Guo and his colleagues (2016) note that LDA–based analysis has been used to examine mass media text, especially when analyzing well–constructed text such as news articles. Based on a collection of documents, a topic model can detect a set of “topics” and assess how strong each topic is exhibited throughout the text. Maier (2018) highlights that LDA topic model is a powerful method that can uncover latent correlations between words, even if they are not presented in a text together.

This study applied the Structural Topic Model (STM), which is a topic model technique that builds on the LDA model. STM shares its features with the LDA model, except it additionally allows for the inclusion of document-level external variables such as particular characteristics of the authors. That is, STM is more powerful than other topic model techniques in that it allows researchers to examine additional information about each document and incorporate such metadata as covariates into the topic model (Roberts et al., 2019). The STM framework postulates that such metadata of the document can influence document-topic proportions and topic-word distributions (Roberts et al., 2014). As this study aims to examine the characteristics of each news media as factors that influence frames and emotions in news stories, STM was the framework most appropriate for the purpose for my research interest.

2. Emotion Analysis

In order to identify discrete emotions used in the news media, this study used two computer-aided emotion analysis methods - LIWC (Pennebaker et al., 2015) and NRC Emotion Lexicon (Mohammad & Turney, 2010). These are two of the most commonly-used computational techniques that are used to identify and analyze discrete emotions embedded in text data. LIWC was used to discover negative emotion – anger, and fear, and NRC Emotion Lexicon to detect both positive and negative emotion – anger, fear and hope.

2.1 LIWC

Linguistic Inquiry and Word Count (LIWC) conducts a word-by-word based analysis of text data and calculates the percentage of words that belong to 74 categories that include emotion words (Bantum & Owen, 2009). It has been frequently used in various researches to analyze text data expressed in media outlets. For example, Rashkin and her colleagues (2017) use LIWC to examine the linguistic characteristics used in fake or untrustworthy news. Moreno et al. (2019) apply LIWC to compare the number of fear-based reporting between news coverage about bullying and cyberbullying. As can be partly inferred from its wide usage, the LIWC program is believed to identify emotional features of text data in an objective and quantitative way (Moreno et al., 2019). In the process of its development, LIWC was validated for its reliability and validity. According to Pennebaker et al. (2001), the interrater reliability discrimination between category word elements ranged from 86% to 100% depending on which dimension was being assessed. Furthermore, when judges rated over 200 essays on different LIWC dimensions, the LIWC and judges' ratings of essays showed moderate to strong level of correlations for majority of the emotion categories, indicating high construct validity (Pennebaker et al., 1997).

2.2 NRC Emotion Lexicon

The National Research Council (NRC) Emotion Lexicon refers to a list of words and emotions that represent a range of distinct emotions that can be used to identify emotions in text data. The NRC Emotion Lexicon analyzes the emotions embedded in text data and categorizes them into eight distinct emotions according to the emotions they represent - anger, anticipation, disgust, fear, joy, sadness, surprise and trust (Mohammad & Turney, 2010). The lexicon was first manually annotated using Amazon's Mechanical Turk service, and was then validated with automatically generated questions to identify and eliminate erroneous annotations. The lexicon was also compared with gold standard data to assess its quality. When the term-emotion annotation results from different annotators were compared, more than 50% of the items had agreements from all annotators and more than 80% of the items had four (out of five) annotators' agreement, implying high level of agreement between coders (Mohammad & Turney, 2010). NRC Emotion Lexicon is also widely used to detect emotions expressed in texts of various media outlets. For instance, Kunmar and his colleagues (2020) use NRC to examine the emotional effects that are caused by social media posts during COVID-19 pandemic crisis. Van den Broek-Altenburg and Ahterly (2019) collect Twitter data about health insurance and apply NRC Emotion Lexicon to examine what sentiments consumers have towards health care providers and health insurance.

With NRC Emotion Lexicon being capable of detecting both positive and negative emotions, I utilized the program to analyze the three targeted emotions – anger, hope and fear. NRC Emotion Lexicon does not have an exact category named "hope", thereby emotions detected as "anticipation" in NRC Emotion Lexicon was coded as "hope" in this paper.

It is necessary to utilize both NRC Emotion Lexicon and LIWC, as doing so will allow for a comparison between the results. By examining whether results yielded from each of the common emotion are consistent with each other, it would be possible to cross– validate the outcome. As LIWC is only capable of detecting negative emotions, the comparison between the two methods was made for fear and anger. It is true that in LIWC, there is no exact emotional category named "fear" – instead, it is capable of detecting "anxiety". For the sake of validity and considering the similarities in the two emotions, I decided to compare the results extracted under the category "fear" in NRC Emotion Lexicon with that of "anxiety" in LIWC.

In order to answer the research questions and hypotheses, this study investigated how the type of media outlet (newspaper vs. TV news) and the political preference of news media organization are linked with the levels of each discrete emotion.

Chapter 5. Results

1. Data collection

Using Nexis Uni, a database that provide access to full-text news articles, I collected news articles published in the US between 2019 and 2021 from the aforementioned nine targeted news outlets (The New York Times, USA Today, The Los Angeles Times, The Wall Street Journal, ABC, CBS, NBC, CNN, FOX). I used "artificial intelligence", "AI", and "A.I." as keywords to retrieve all articles related to artificial intelligence. A total of 5,587 artificial intelligence- focused articles were collected. Table 1 shows the number of news articles retrieved by respective news outlets.

Table 1. Number of news articles collected by news outlets

News media	2019	2020	2021	Total
The New York Times	1423	983	959	3,365
USA today	72	63	61	196
The Los Angeles Times	217	101	164	482
The Wall Street Journal	114	104	98	316
ABC	28	29	35	92
CBS	49	33	35	117
NBC	57	19	35	111
CNN	314	187	212	713
FOX	91	43	61	195
TOTAL	2365	1562	1660	5587

Although the number of articles collected each year (2019, 2020, 2021) showed little variance, there was a prominent difference in the number of articles collected by each media outlet. In particular, the number was significantly high for The New York Times.

In terms of the variables of interest, it is worth mentioning that the numbers were imbalanced. 4087 articles were collected from 솟 liberal media and 511 from the conservative media. There were 4,359 articles collected from newspapers and 1,228 from TV news. It is noteworthy that there were stark differences – media outlets with liberal political orientation and in the form of newspapers were more dominant than their counterparts.

2. Frames

Prior to analysis, the text data from news articles were processed in order to suit the preprocessing standards of topic modeling. I converted the letters to lowercase, removed numbers, punctuations and stop words, and stemmed the data by facilitating STM R package.

Next, I selected the optimal number of topics the STM can yield. The structural topic model technique has been commonly criticized that the results can be influenced by how the parameters are set, especially by the number of identified topics (Maier et al., 2018). So as to tackle this problem, researchers commonly run multiple models using different number of topics and compare the results (Maier et al., 2018).

As previous researches point out (Grimmer and Stewrat, 2013; Roberts et al., 2019), there is no 'right' answer when choosing the most appropriate number of topics of a given corpus. Rather, individual researchers have the burden to decide what the optimal number is, based on the context or several automated tests the STM provides such as exclusivity, held-out likelihood, residuals, or semantic coherence. Amongst these quantities of interest, one widely-used method is to compare average exclusivity and semantic coherence level. The level of semantic coherence gets higher when the most probable words for a particular topic co-occur frequently. Hence, this metric often goes in accordance with human judgement of topic quality. However, high semantic coherence can be relatively easy to have, as a few topics with very common words can maximize this level. One solution to this problem is to observe exclusivity of words to topics. Semantic coherence and exclusivity are naturally in a negative relationship - the higher the semantic coherence, the lower the exclusivity.

Following such logic, I compared semantic coherence and exclusivity retrieved when different values for K (the number of topics) are applied (Roberts et al., 2014;2019), setting the range of K from 10 to 100. The result showed optimal balance when the number of topics was 20 or 30, but I chose 20, given that the model was contextually more interpretable than the one with 30 topics.

RQ1 aims to observe the salient frames in news coverage of artificial intelligence in the United States. As aforementioned, I identified a total of twenty frames, using structural topic modeling to collect words with high probability and FREX (frequency exclusivity) scores. While words collected according to probability scores are those that appear most commonly, the ones collected according to FREX scores show high frequency and exclusivity. Therefore, by considering both probability and FREX scores, I aimed to obtain a more comprehensive and appropriate understanding.

After extracting the topics, each of them was labeled, based on common themes or characteristics that can represent the extracted words. Table 2 shows the topics retrieved from the nine targeted media outlets. The twenty topics were labeled as the following – "future challenges (topic 1)", "data bias (topic 2)", "health (topic 3)", "job (topic 4)", "governance and regulations (topic 5)", "politics and the President (topic 6)", "photos and videos (topic 7)", "art and culture (topic 8)", "technological research and development (topic 9)", "general words (topic 10)", "industrial interest (topic 11)", "national security and global economy (topic 12)", "general words (topic 13)", "social media and media platforms (topic 14)", "business and economy (topic 15)", "general words (topic 16)", "general words (topic 17)" "economic prospects (topic 18)" "general words (topic 19)" and "industrial interest II (topic 20)".

Topic 1 included words such as 'challenge', 'respond', 'question', or 'action', which were about the future issues or challenges related to AI that is left to be addressed – hence given the label "future challenges". Topic 5 consisted words such as 'law', 'govern', 'rule', or 'state', which boil down to the issues of how to regulate or govern the new technology, therefore was titled "governance and regulations". Topic 7 was comprised of words such as 'photo', digital', 'image', video', face', or 'create'. It could be inferred that these words refer to the issue of AI technology being used to create images or video footages of a fake person or replace an image/video of a person with someone else's image. Topic 7 was, therefore, named ‘photos and videos’. Topic 8 included words such as ‘culture’, ‘book’ write’, ‘sense’ or ‘imagine’ which relate to AI writing books and creating art, raising questions of whether AI is capable of producing and appreciating art or culture which demand imagination or feelings. Topic 8 was, hence, titled “art and culture”. The rest of the topic labels are self-explanatory.

Topic 10, 13, 16, 19 were labeled as “general words”, consisting a collection of words that showed little consistency or common characteristics. This is due to the fact that in STM, topics are generated based on the distribution over words without considering the meanings of the words. That is, words that repeatedly occur together are automatically generated into a topic, leading to topics without prominent themes. In this case, topic 10,13, 16, and 19 fell into this category.

Table 2. Topics of news media

Topic no.	Label	Criteria	Words
1	Future challenges	Probability	need, challeng, world, respons, meet, respond, question, leader, action, make, issu, commit, clear, whether, support, communiti, address, must, face, concern
		FREX	challeng, respond, respons, meet, action, commit, leader, address, must, clear, need, question, togeth, communiti, support, threat, engag, intern, decis, world
2	Data bias	Probability	polic, view, peopl, citi, white, offic, black, attack, fire, use, law, depart, one, kill, charg, man, group, communiti, face, street
		FREX	polic, view, citi, black, attack, white, fire, offic, kill, depart, charg, man, peopl, law,

			street, communiti, local, stop, imag, face
3	Health	Probability	case, test, health, peopl, state, answer, pandem, number, countri, offici, death, get, public, govern, spread, say, safeti, effect, risk, expert
		FREX	case, test, answer, health, death, pandem, number, spread, safeti, offici, risk, safe, expert, effect, state, countri, author, measur, die, posit
4	Job	Probability	work, job, school, care, worker, women, american, peopl, children, famili, pay, need, letter, home, help, health, mani, parent, make, get
		FREX	school, job, care, women, worker, letter, children, famili, work, pay, american, parent, home, health, benefit, america, age, program, high, need
5	Governance and regulations	Probability	data, technolog, use, law, inform, compani, govern, articl, regul, agenc, public, protect, collect, system, feder, provid, person, servic, rule, state
		FREX	data, law, articl, inform, regul, agenc, technolog, protect, collect, govern, feder, use, rule, requir, public, provid, identifi, internet, track, servic
6	Politics and the President	Probability	presid, trump, democrat, biden, hous, senat, elect, vote, polit, campaign, state, white, american, bill, parti, former, hes, want, support, nation
		FREX	democrat, presid, biden, senat, trump, vote, hous, elect, campaign, polit, parti, bill, hes, opinion, debat, congress, white, former, win, washington
7	Photos and videos	Probability	use, app, photo, peopl, graphic, phone, camera, digit, make, imag, video, user, person, onlin, tool, like, also, help, face, creat

		FREX	app, graphic, phone, camera, photo, digit, imag, tool, use, user, voic, onlin, track, connect, softwar, avail, person, free, creat, video
8	Art and culture	Probability	one, book, life, world, stori, way, like, year, write, live, read, peopl, even, power, human, seem, cultur, becom, histori, time
		FREX	book, life, stori, write, read, cultur, histori, seem, idea, live, experi, word, feel, becom, way, might, sens, war, imagin, age
9	Technological research and development	Probability	research, system, technolog, human, comput, use, univers, machin, learn, develop, work, like, advanc, scienc, studi, one, help, program, engin, way
		FREX	research, comput, machin, system, advanc, learn, human, scienc, univers, develop, studi, train, professor, technolog, predict, engin, program, field, design, understand
10	General words	Probability	report, investig, note, releas, potenti, say, general, press, comment, accord, involv, call, inform, suggest, use, statement, target, sourc, help, find
		FREX	report, note, investig, releas, press, potenti, comment, involv, general, suggest, sourc, accord, statement, target, find, cover, inform, say, activ, order

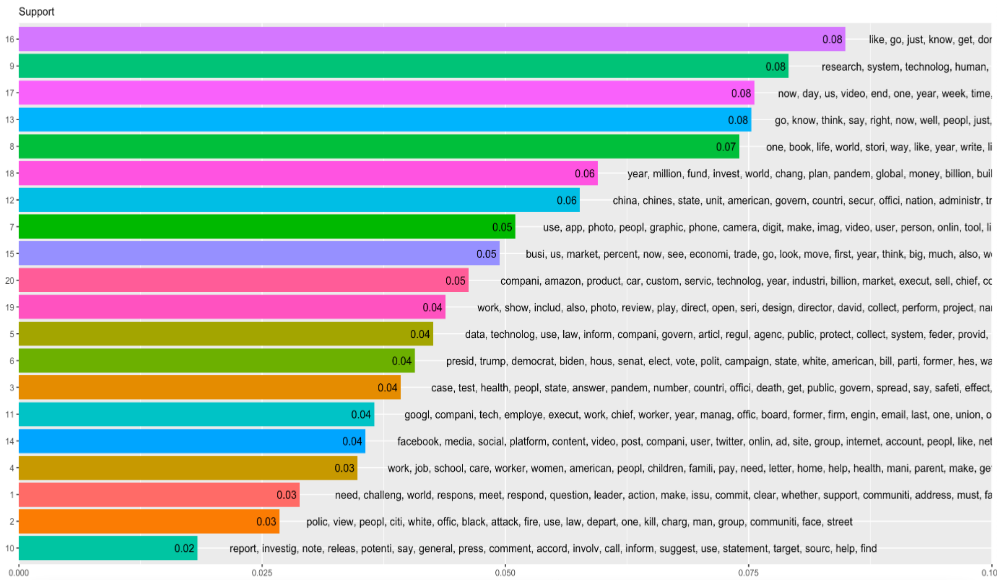
11	Industrial interest	Probability	googl, compani, tech, employe, execut, work, chief, worker, year, manag, offic, board, former, firm, engin, email, last, one, union, organ
		FREX	googl, tech, employe, execut, compani, chief, manag, board, worker, engin, firm, email, former, union, offic, team, organ, work, project, wrote
12	National security and global economy	Probability	china, chines, state, unit, american, govern, countri, secur, offici, nation, administr, trade, us, militari, trump, foreign, war, world, econom, technolog

		FREX	china, chines, unit, administr, secur, american, militari, foreign, offici, trade, state, countri, govern, war, econom, nation, europ, global, deal, america
13	General words	Probability	go, know, think, say, right, now, well, peopl, just, that, yes, get, us, thank, clip, see, look, realli, want, come
		FREX	know, yes, go, think, thank, that, clip, well, right, realli, theyr, mean, talk, lot, actual, just, now, im, your, say
14	Social media and media platforms	Probability	facebook, media, social, platform, content, video, post, compani, user, twitter, onlin, ad, site, group, internet, account, peopl, like, network, mark
		FREX	facebook, platform, content, media, social, post, twitter, user, video, onlin, site, internet, ad, account, mark, group, network, spread, polici, children
15	Business and economy	Probability	busi, us, market, percent, now, see, economi, trade, go, look, move, first, year, think, big, much, also, week, well, deal
		FREX	busi, market, percent, economi, trade, move, rate, see, cours, price, us, weve, big, deal, expect, term, econom, now, look, cut
16	General words	Probability	like, go, just, know, get, dont, im, that, want, think, thing, one, right, good, your, got, game, say, realli, time
		FREX	im, dont, game, know, that, your, got, just, love, thing, go, good, guy, didnt, like, realli, feel, ive, littl, lot
17	General words	Probability	now, day, us, video, end, one, year, week, time, back, begin, first, say, still, two, live, see, last, hour, come
		FREX	day, end, begin, hour, week, morn, featur, video, now, back, still, night, watch, team, live, th, next, home, latest, us
18	Economic prospects	Probability	year, million, fund, invest, world, chang, plan, pandem, global, money, billion, build, last, like, partner, space, travel, mani, one, need

		FREX	fund, invest, million, partner, money, pandem, global, plan, travel, space, capit, build, billion, year, chang, near, crisi, area, manag, spend
19	General words	Probability	work, show, includ, also, photo, review, play, direct, open, seri, design, director, david, collect, perform, project, name, origin, whose, interview
		FREX	review, show, seri, direct, david, play, perform, origin, open, design, director, piec, whose, includ, collect, present, photo, event, project, work
20	Industrial interest II	Probability	compani, amazon, product, car, custom, servic, technolog, year, industri, billion, market, execut, sell, chief, consum, last, like, offer, tech, make
		FREX	amazon, custom, car, product, compani, servic, sell, industri, billion, consum, market, buy, chief, offer, giant, firm, execut, technolog, drive, oper

In order to address RQ1, I organized the topics by prevalence (see Figure 1 below).

Figure 1. Topics by prevalence



After eliminating topics labeled "general topics" which have no relationships with artificial intelligence−related discourses, "technological research and development (topic 9)" showed the highest level of prevalence. The next was followed by "art and culture (topic 8)", "economic prospects (topic 18)", "national security and global economy (topic 12)", and "photos and videos (topic 7)".

Although extracted as separate topics, there were five topics (topic 11,12,15,18,20) that were related to economy – "industrial interest (topic 11)", "national security and global economy (topic 12)", "business and economy (topic 15)", "economic prospects (topic 18)" and "industrial interest II (topic20)". Considering the fact that the total number of topics extracted was 20, it is noticeable that economy−related topics took up 25% of the total. This indicates that economic frames have high salience in the discourse around artificial intelligence.

Amongst the least prominent topics were "jobs (topic 4)", "future challenges (topic1)" and "data bias (topic2)". "Jobs (topic 4)" consisted words about whether jobs and educational system maintained by human resources can be/ will be substituted by artificial intelligence. "Future challenges (topic 1)" was about the concerns and challenges related to AI that need to be addressed. "Data bias (topic2)" posed that the data utilized by AI may reflect existing discriminatory bias on issues such as gender or race. The topics with low level of prevalence shared a common grounding in that they show concerns about the changes artificial intelligence may bring or highlight ethical or social issues that need to be addressed.

So as to examine H1 and H2, I compared each topic with the political ideology of media outlets and the type of media. The following table shows the regression for each topic.

Table 3. Regression of topic prevalence in news media

	Topic1	Topic2	Topic3	Topic4	Topic5
Intercept	0.003*** (0.001)	0.029*** (0.001)	0.043*** (0.001)	0.026*** (0.001)	0.006*** (0.001)
Newspaper (vs. TV news)	0.001* (0.001)	−0.005*** (0.001)	−0.005*** (0.001)	0.011*** (0.001)	0.043*** (0.001)
Conservative (vs. liberal)	−0.011*** (0.001)	−0.006** (0.002)	−0.003 (0.002)	−0.000 (0.002)	0.035*** (0.002)
Neutral (vs. liberal)	0.000 (0.000)	0.011*** (0.001)	0.001 (0.001)	−0.001 (0.001)	−0.005** (0.002)
	Topic6	Topic7	Topic8	Topic9	Topic10
Intercept	0.074*** (0.001)	−0.001 (0.002)	0.044*** (0.002)	0.018*** (0.003)	0.012*** (0.001)
Newspaper (vs. TV news)	−0.040*** (0.001)	0.053*** (0.002)	0.045*** (0.002)	0.077*** (0.003)	0.006*** (0.001)
Conservative (vs. liberal)	0.012*** (0.002)	0.057*** (0.003)	−0.037*** (0.004)	0.019*** (0.004)	0.013*** (0.0007)
Neutral (vs. liberal)	−0.022*** (0.002)	0.027*** (0.002)	−0.011*** (0.003)	−0.006. (0.003)	0.001 (0.001)
	Topic11	Topic12	Topic13	Topic14	Topic15
Intercept	0.006** (0.002)	0.041*** (0.003)	0.307*** (0.001)	0.016*** (0.002)	0.075*** (0.001)
Newspaper (vs. TV news)	0.038*** (0.002)	0.028*** (0.003)	−0.292*** (0.001)	0.024*** (0.002)	−0.025*** (0.001)
Conservative (vs. liberal)	0.002 (0.003)	−0.013** (0.005)	0.021*** (0.002)	−0.015*** (0.003)	−0.024*** (0.002)
Neutral (vs. liberal)	−0.001 (0.002)	−0.023*** (0.003)	−0.040*** (0.001)	0.008** (0.002)	0.022*** (0.002)
	Topic16	Topic17	Topic18	Topic19	Topic20
Intercept	0.104*** (0.003)	0.110*** (0.002)	0.036*** (0.002)	0.009*** (0.001)	0.008*** (0.002)
Newspaper (vs. TV news)	−0.039*** (0.003)	−0.042*** (0.002)	0.032*** (0.002)	0.043*** (0.001)	0.043*** (0.002)

Conservative (vs. liberal)	0.003 (0.004)	−0.004*** (0.003)	−0.015*** (0.003)	−0.006** (0.002)	0.014*** (0.003)
Neutral (vs. liberal)	0.056*** (0.003)	0.012*** (0.002)	−0.004. (0.002)	0.007*** (0.001)	0.011*** (0.002)

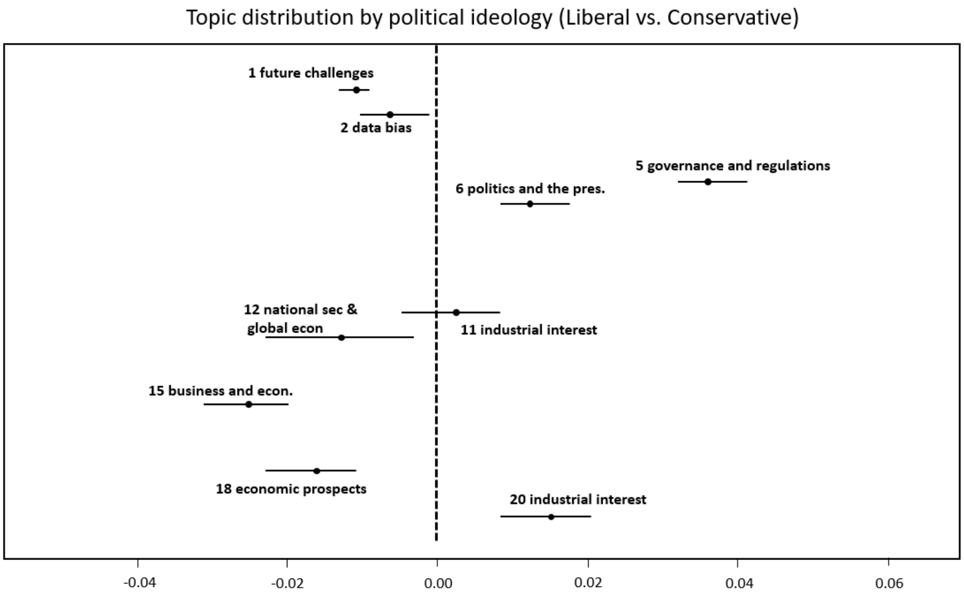
*** p<.001, ** p<.01, *p<.05, The values in parentheses are standard errors

In order to address H1a and H1b, I compared topics by political orientation of media outlets. H1a predicted that conservative media will more likely discuss AI using frames about economic benefits, whereas H1b predicted that liberal media will more likely discuss preventive measures to tackle issues that can be caused by AI.

Figure 2 displays the distribution of topics by the political ideology of media outlets (Figure 2 only includes topics related to the hypotheses. See appendix to see distribution of all topics). As the distribution shows, the topics that focus on economic aspects of artificial intelligence-related discourse such as "national security and global economy (topic 12)", "business and economy (topic 15)" and "economic prospects (topic 18)" were more likely to be spotted in the liberal media, rejecting H1a. However, it was noteworthy that the two topics named "industrial interest (topic 11 and topic 20) were more likely to be expressed in conservative media. This indicates that although grouped under the category of 'economy' in a broad sense, there were differences stemming from which discourse each topic focused on in detail.

H1b was also not fully supported. It is true that topics like "future challenges (topic1)"and "data bias (topic 2)" which pose challenges and issues around AI that remains to be tackled were more prevalent in liberal media outlets. On the other hand, topics such as "governance and regulations (topic 5)" that deal with measures to regulate AI turned out to be more likely to be found in conservative media. There was no clear tendency found in terms of the relations between frames and political orientation.

Figure 2. Topic distribution by political ideology

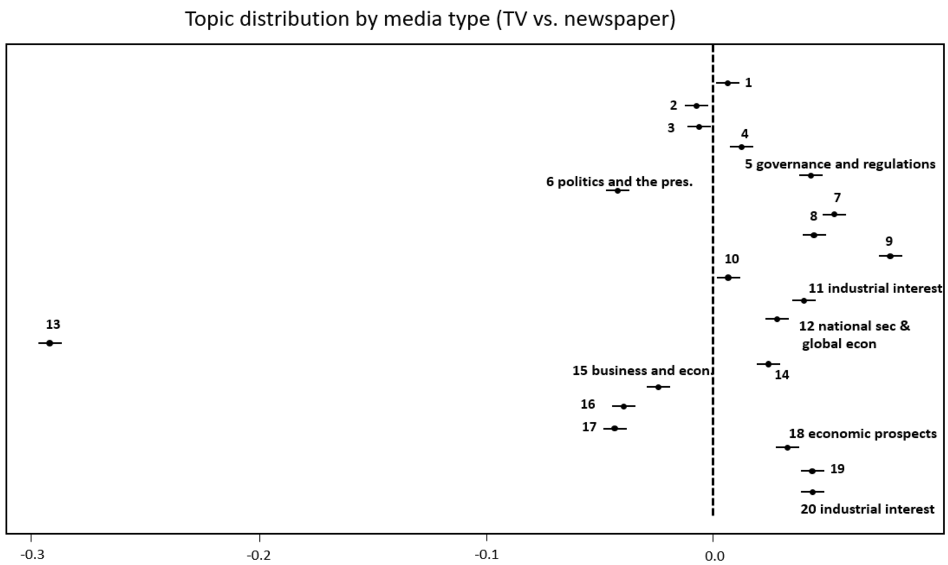


RQ2 explored whether there were meaningful differences in the frequently used frames in news coverage about AI between newspaper and TV news. H2 predicted that newspapers will be more likely than TV news to use frames discussing economic discourses about AI.

Figure 3 displays the distribution of topics by the type of media. As can be observed, topics related to economic frames – “industrial interest (topic 11)”, “national security and global economy (topic 12)”, “economic prospects (topic 18)” and “industrial interest (topic 20)” were more likely to be observed in newspapers than TV news as hypothesized (H2).

However, it is worth mentioning that the distribution of topics leaned towards newspapers in general. After eliminating topics labeled "general words", it could be observed that "politics and the president (topic6)", "business and economy (topic 15)", "health (topic3)" and "data bias (topic2)" were the only topics that were more likely to be spotted in TV news rather than newspapers. Considering that the value for "health (topic 3)" (*B=−0.005, SE=0.001, p<0.001)* and "data bias (topic2)" (*B=−0.005, SE=0.001, p<0.001)* were relatively low, the most prominent topics related to TV news were "politics and the president (topic6)" and "business and economy (topic15)". The rest of the topics were more likely to be found in newspapers. This indicates that while the results support the hypothesis that newspapers are more likely to discuss economic discourses, such tendency is not only limited to economy−related discourses.

Figure 3. Topic distribution by media type



3. Emotion Analysis

RQ3 addressed the question of whether there are meaningful differences in the discrete emotions expressed in the news about AI depending on the political ideology of each media outlet. H3a predicted that conservative media will be more likely to use hope and liberal media will be more likely to use fear and anger in AI related articles. H3b predicted that liberal media will be more likely than conservative media to use fear and anger.

RQ4 investigated the question of whether there are meaningful differences in the discrete emotions expressed in the news media about AI depending on the type of media outlet. H4a posed that TV news will be more likely than newspapers to use fear and anger, while H4b predicted that newspapers will be more likely than TV news to use hope to talk about AI.

In order to examine the said hypotheses, I used NRC Emotion Lexicon and LIWC to detect emotions in each article.

3.1. NRC Emotion Lexicon

Table 4. Regression on level of emotion (NRC Emotion Lexicon)

	Emotion		
Predictor	Anger	Hope	Fear
(intercept)	0.112*** (0.046)	0.216*** (0.051)	0.172*** (0.063)

Conservative (vs. Liberal)	0.076* (0.037)	0.310*** (0.041)	0.014 (0.051)
Neutral (vs. Liberal)	0.042 (0.027)	0.123*** (0.030)	−0.042 (0.037)
Newspaper (vs. TV news)	−0.034 (0.038)	−0.072. (0.042)	−0.016 (0.052)
Number of words	0.000 ** (0.000)	−0.000 (0.000)	0.000* (0.000)

*** p<.001, ** p<.01, *p<.05. The values in parentheses are standard errors

To investigate RQ3, I observed the relationship between the level of emotions and the political orientation of media outlets, controlling for message length (number of words). As can be observed in Table 4, the level of hope (coded as ‘anticipation’ in NRC Emotion Lexicon) showed positive relationship with media outlets with conservative political orientation (b = 0.310, SE_b = 0.041, p < .001). This result demonstrates that conservative media are more likely than the liberal media to use hope to discuss AI related articles, as predicted by H3a. In regards to H3b, the result showed that anger (b = 0.076, SE_b = 0.037, p < .05) showed positive correlation with conservative political orientation, albeit weak. The political ideology of media outlets was statistically not significant for fear (b = 0.014, SE_b = 0.051, p >.05). Accordingly, H3b was rejected.

In order to investigate RQ4, I examined the correlations between emotions and type of media. Table 4 shows that the type of media was statistically not significant for the level of of anger (b = −0.134, SE_b = 0.038, p >.05), hope (b = −0.072, SE_b = 0.042, p >.05) and fear (b = −0.016, SE_b = 0.052, p >.05). This indicates that none of the three emotions were affected by the type of media used.

3.2. LIWC

In order to cross-validate the results extracted by NRC Emotion Lexicon, I used LIWC to detect the targeted discrete emotions. As mentioned earlier, LIWC is capable of examining only the negative emotions. Therefore, I used LWC to examine the emotion of anger and fear (coded by LIWC as 'anxiety') and compared the results with that of NRC Emotion Lexicon.

Table 5. Regression on level of emotion (LIWC)

	Anger	Anxiety
(Intercept)	0.131*** (0.012)	0.101*** (0.006)
Conservative (vs. Liberal)	0.070*** (0.014)	0.051*** (0.014)
Neutral (vs. Liberal)	0.013. (0.007)	−0.0006 (0.007)
Newspapers (vs. TV)	0.016. (0.010)	0.005 (0.009)
Number of words	−0.000*** (0.000)	−0.000*** 0.000)

*** p<.001, ** p<.01, *p<.05, The values in parentheses are standard errors

To explore RQ3, regression coefficients estimated as can be seen in Table 5. The results indicate that anger (b = 0.070, SE_b = 0.014, $p < .001$) and anxiety (b = 0.051, SE_b = 0.014, $p < .001$) were displayed greater in conservative media as opposed to the liberal, going against H3b. This outcome was loosely consistent with that of NRC Emotion Lexicon. The data from both NRC Emotion Lexicon and LIWC contended that anger showed positive relation with conservatism. In terms of fear, while it is true that the correlation was statistically not significant in NRC Emotion Lexicon, the numbers still indicate that the level of fear was in positive relationship with conservative orientation (b = 0.014, SE_b = 0.051, p >.05). Considering the fact that fear also showed positive correlation with conservatism in LIWC, it can be said that the two methods loosely share the same tendency.

In terms of RQ4, Table 5 shows that the type of media is not statistically significant for both anger (b = 0.016, SE_b = 0.010, p >.05) and anxiety (b = 0.005, SE_b = 0.009, p >.05). This means that the two emotions are not influenced by which type of media is used, thereby rejecting H4a. However, this result was consistent with the one from NRC Emotion Lexicon.

Chapter 6. Discussion

Through topic modeling, this paper made attempts to detect frames expressed in news articles and identify their relationship with the political orientation of media outlets and the type of media. The data showed that economy-related discourse was the most prevalent frame used to discuss artificial intelligence. However, the analytic results did not correspond with the hypotheses – that is, there was no clear inclination observed in the correlations between the extracted topics and the political orientation.

It could be inferred that such outcome may be due to the fact that AI is still in the early stage of the issue cycle. As artificial intelligence technology is in the primary level of development, it has only recently entered the realm of public discourse. Accordingly, news coverage about AI may be largely focused on filling in information deficits regardless of political viewpoints. That is, discussions around AI may be primarily focused on technological possibilities, and have not yet entered the realm of discussing social implications that lead to political differences. As the discussions around AI is still in flux, the inclination each media aims to portray is not fixed yet, leading to mixed distribution of topics.

It is also noteworthy that AI-related articles were more prevalent in newspapers rather than TV news in general. Such outcome may result from the complexity and depth of knowledge involved in discussing technological issues like artificial intelligence. The scientific, political, social, and economic aspects of AI technology are all likely to entail detailed explanation about background information or context. Therefore, it is reasonable to conclude that subjects like AI are more suitable to be discussed in newspapers.

In terms of the emotional analysis, one major finding is that emotion was not influenced by the type of media used. That is, the level of anger, hope and fear were indifferent in both newspapers and

TV news. It can be said that this finding goes in line with the implication of the STM analysis – that is, as AI is still in the early stage of the issue cycle, both TV and newspapers may be focused on providing information, relatively detached from the emotions.

While comparing the outcomes from NRC Emotion Lexicon and LIWC, I found that the two show similar tendency in general. However, there still existed minor differences – while NRC Emotion showed that fear showed no statistically-significant correlation with conservative orientation, LIWC showed that anxiety and conservatism hold positive relations. One possible explanation for such outcome may be that, while coded as the same in this paper, the emotion of 'fear' and 'anxiety' are not perfect equivalents. Hence, one limitation of this paper is that for the sake of guaranteeing validity, it regarded two similar – but different- emotions as equivalent.

Another possible reason may stem from the difference in the two computational analytic methods. In attempts to interpret th differing results between the two methods, I took closer look at the data. I noticed that the N value (the number of articles that express anger) for NRC Emotion Lexicon was 5,146, whereas for LIWC it was only 3,187. In addition, the level of anger detected in news articles also showed difference depending on the analytic tool utilized. While the level averaged around 1.1592 for NRC Emotion Lexicon, he numbers were significantly low for LIWC, averaging only 0.1293. (See Table II in the appendix). For fear, the average was 1.754 for NRC Emotion Lexicon, and 0.1336 for LIWC. Through such comparison, I noticed that there were clear differences in how and how sensitive the two tools detected emotion. Such differences may explain for why the two methods yielded different outcome.

One solution to prevent such confusion could be the involvement of human coders. Validity has been one of the most highly contested issues regarding computer-aided analytic methods – whether computers can adequately detect frames or emotions that are actually present in a certain text. Multiple researchers suggest the use of human intervention to make it check whether the results match adequate interpretation of text documents (Walter and Ophir, 2019; Guo et al., 2016; Maier, 2018). That is, human coders can be made to detect frames or emotions in a random sample of news stories used for computational analysis. By comparing the results retrieved by computers and human coders, it would be possible to verify whether the computational methods successfully analyzed the given data. In future studies, the confusion met by this paper could be overcome by utilizing human intervention.

One other limitation of this paper is that it only collected news articles from 2019 to 2021. The three years' worth of articles may have not been enough to provide a holistic and balanced understanding of the overall inclination of discourses around artificial intelligence. Based on the outcome retrieved from this paper, I plan to collect data from longer time period in future studies.

Appendix

Figure I. Topic distribution by political ideology

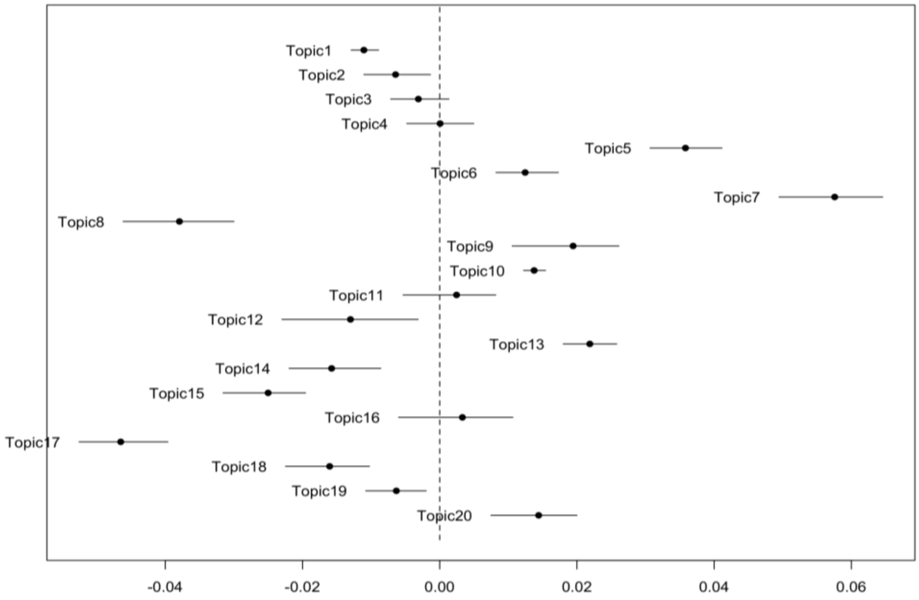
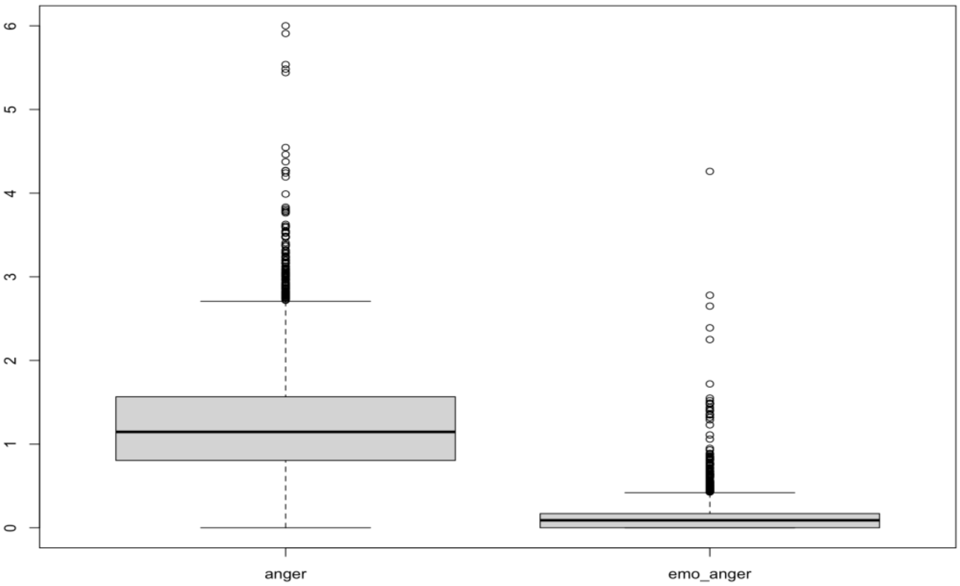
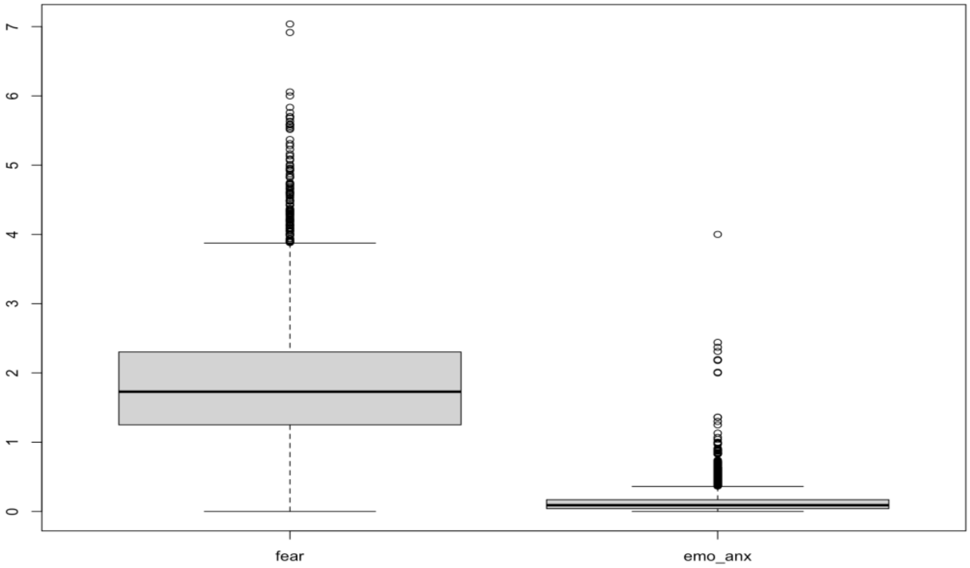


Table I. Box plots of level of emotion extracted by NRC Emotion Lexicon and LIWC





* left – NRC Emotion Lexicon / right – LIWC

Bibliography

Anderson, A. (2014). *Media, environment and the network society*. Basingstoke: Palgrave Macmillan.

Bantum, E. O. C., & Owen, J. E. (2009). Evaluating the validity of computerized content analysis programs for identification of emotional expression in cancer narratives. *Psychological assessment*, *21*(1), 79.

Burscher, B., Odijk, D., Vliegenthart, R., Rijke, & Vreese, C. (2014) Teaching the computer to code frames in news: Comparing two supervised machine learning approaches to frame analysis. *Communication Methods and Measures*, 8(3), 190-206.

Budak, C., Goel, S., & Rao, J. M. (2016). Fair and balanced? Quantifying media bias through crowdsourced content analysis. *Public Opinion Quarterly*, *80*(S1), 250-271.

Caughey, D., & Warshaw, C. (2018). Policy preferences and policy change: Dynamic responsiveness in the American states, 1936–2014.

Cobb, M. D. (2005). Framing effects on public opinion about nanotechnology. *Science communication*, *27*(2), 221-239.

Cooper, C. P., & Roter, D. L. (2000). " If it bleeds it leads"? Attributes of TV health news stories that drive viewer attention. *Public Health Reports*, *115*(4), 331.

Cho, J., Boyle, M. P., Keum, H., Shevy, M. D., McLeod, D. M., Shah, D. V., & Pan, Z. (2003). Media, terrorism, and emotionality: Emotional differences in media content and public reactions to the September 11th terrorist attacks. *Journal of Broadcasting & Electronic Media*, *47*(3), 309-327.

De los Santos, T. M., & Nabi, R. L. (2019). Emotionally charged: Exploring the role of emotion in online news information seeking and processing. *Journal of Broadcasting & Electronic Media*, *63*(1), 39-58.

Driedger, S. M. (2007). Risk and the media: a comparison of print and televised news stories of a Canadian drinking water risk event. *Risk Analysis: An International Journal*, *27*(3), 775-786.

Druckman, J. N. (2005). Media matter: How newspapers and

television news cover campaigns and influence voters. *Political communication*, *22*(4), 463-481.
Entman, R.M. (1993). Framing: Toward clarification of a fractured paradigm. Journal of Communication, 43(3), 51-58.
Eveland Jr, W. P., Seo, M., & Marton, K. (2002). Learning from the news in campaign 2000: An experimental comparison of TV news, newspapers, and online news. *Media Psychology*, *4*(4), 353-378.
Fiske S.T. & S.E. Taylor, 1991. Social Cognition (2nd ed.). New York: McGraw-Hill
Flaxman, S., Goel, S., & Rao, J. M. (2016). Filter bubbles, echo chambers, and online news consumption. *Public opinion quarterly*, *80*(S1), 298-320.
Gamson, W. A., & Modigliani, A. (1994). The changing culture of affirmative action. *Equal employment opportunity: labor market discrimination and public policy*, *3*, 373-394.
Garvey, C., & Maskal, C. (2020). Sentiment analysis of the news media on artificial intelligence does not support claims of negative bias against artificial intelligence. *Omics: a journal of integrative biology*, *24*(5), 286-299.
Gaskell, G., Bauer, M. W., Durant, J., & Allum, N. C. (1999). Worlds apart? The reception of genetically modified foods in Europe and the US. *Science*, *285*(5426), 384-387.
Gross, K. (2008). Framing persuasive appeals: Episodic and thematic framing, emotional response, and policy opinion. Political Psychology, 29, 169-192.
Guo, L., Vargo, C. J., Pan, Z., Ding, W., & Ishwar, P. (2016). Big social data analytics in journalism and mass communication: Comparing dictionary-based text analysis and unsupervised topic modeling. *Journalism & Mass Communication Quarterly*, *93*(2), 332-359.
Holden, C. 2004. California's proposition 71 launches stem cell gold rush. Science 306(5699): 1111.
Izard, C. E. 1977. *Human emotions*, New York, NY: Plenum Press.
Iyengar, S. (1994). *Is anyone responsible?: How television frames political issues*. University of Chicago Press.
Kim, H. J., & Cameron, G. T. (2011). Emotions matter in crisis: The role of anger and sadness in the publics' response to crisis

news framing and corporate crisis response. *Communication Research*, *38*(6), 826–855.

Kim, S. H., Carvalho, J. P., & Davis, A. C. (2010). Talking about poverty: News framing of who is responsible for causing and fixing the problem. *Journalism & Mass Communication Quarterly*, *87*(3–4), 563–581.

Kühne, R., & Schemer, C. (2015). The emotional effects of news frames on information processing and opinion formation. *Communication Research*, *42*(3), 387–407.

Kwon, K. H., Chadha, M., & Wang, F. (2019). Proximity and networked news public: Structural topic modeling of global Twitter conversations about the 2017 Quebec mosque shooting. *International Journal of Communication*, *13*, 24.

Kunmar, A., Khan, S.U., Kalra A. (2020). COVID–19 pandemic: A sentiment analysis: A short review of the emotional effects produced by social media posts during this global crisis. European Heart Journal, 41(39), 3782–3783.

Lazarus, R.S. (1991). *Emotion and adaptation*. New York, NY: Oxford University Press.

Lee, C. J., Scheufele, D. A., & Lewenstein, B. V. (2005). Public attitudes toward emerging technologies: Examining the interactive effects of cognitions and affect on public attitudes toward nanotechnology. *Science communication*, *27*(2), 240–267.

Merenkov, A. V., Campa, R., & Dronishinets, N. P. (2021). Public opinion on artificial intelligence development. *KnE Social Sciences*, 565–574.

Miller, J. D. (1998). The measurement of civic scientific literacy. *Public understanding of science*, *7*, 203–223.

Mohammad, S., & Turney, P. (2010). Emotions evoked by common words and phrases: Using mechanical turk to create an emotion lexicon. In *Proceedings of the NAACL HLT 2010 workshop on computational approaches to analysis and generation of emotion in text* (pp. 26–34).

Moreno, M. A., Gower, A. D., Brittain, H., & Vaillancourt, T. (2019). Applying natural language processing to evaluate news media coverage of bullying and cyberbullying. *Prevention science*, *20*(8), 1274–1283.

Nabi, R. L. (1999). A cognitive-functional model for the effects of discrete negative emotions on information processing, attitude change, and recall. *Communication theory*, *9*(3), 292−320.
Nabi, R. (2002). Anger, fear, uncertainty, and attitudes: A test of the cognitive−functional model. *Communication Monographs*, *69*(3), 204−216.
Nabi, R. L. (2003). Exploring the framing effects of emotion: Do discrete emotions differentially influence information accessibility, information seeking, and policy preference?. *Communication Research*, *30*(2), 224−247.
Nabi, R. L. (2007). Emotion and persuasion: A social−cognitive perspective. *Communication and social cognition: Theories and methods*, 377−398.
Nabi, R. L. (2010). The case for emphasizing discrete emotions in communication research. *Communication Monographs*, *77*(2), 153−159.
Nisbet, M. C., Brossard, D., & Kroepsch, A. (2003). Framing science: The stem cell controversy in an age of press/politics. *Harvard International Journal of Press/Politics*, *8*(2), 36−70.
OECD (2019). Artificial Intelligence in Society. Paris: OECD Publishing.
Page, B., and R. Shaprio. 1992.The rational public. Chicago: The University of Chicago Press.
Pennebaker, J. W., Francis, M. F., & Booth, R. J. (2001). *Linguistic inquiry and word count: LIWC−2001 operations manual*. Austin, TX: LIWC.net
Pennebaker, J. W., Mayne, T. J., & Francis, M. E. (1997). Linguistic predictors of adaptive bereavement. *Journal of Personality and Social Psychology*, 72(4), 863- 871.
Popkin S.L., 1994 The Reasoning Voter: Communication and Persuasion in Presidential Campaigns (2nd ed.). Chicago, IL: University of Chicago Press.
Potter, W. J., & Riddle, K. (2007). A content analysis of the media effects literature. *Journalism & Mass Communication Quarterly*, *84*(1), 90−104.
Rashkin, H., Choi, E., Jang, J. Y., Volkova, S., & Choi, Y. (2017, September). Truth of varying shades: Analyzing language in fake news and political fact−checking. In *Proceedings of the*

2017 conference on empirical methods in natural language processing (pp. 2931−2937).
Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., ... & Rand, D. G. (2014). Structural topic models for open-ended survey responses. *American Journal of Political Science*, *58*(4), 1064−1082.
Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). Stm: An R package for structural topic models. *Journal of Statistical Software*, *91*(1), 1−40.
Roseman, I. J. (1984). Cognitive determinants of emotion: A structural theory. Review of Personality & Social Psychology, 5, 11–36.
Smith, N., & Leiserowitz, A. (2014). The role of emotion in global warming policy support and opposition. Risk Analysis, 34, 937−948.
Scheufele, D. A. (1999). Framing as a theory of media effects. *Journal of communication*, *49*(1), 103−122.
Scheufele, D. A., & Lewenstein, B. V. (2005). The public and nanotechnology: How citizens make sense of emerging technologies. *Journal of Nanoparticle Research*, *7*(6), 659−667.
Scheufele, D. A., & Tewksbury, D. (2007). Framing, agenda setting, and priming: The evolution of three media effects models. *Journal of communication*, *57*(1), 9−20.
Walter, D., & Ophir, Y. (2019). News frame analysis: an inductive mixed−method computational approach. *Communication Methods and Measures*, *13*(4), 248−266.
Wanta, W. (1997). The messenger and the message: Differences across media. In M. McCombs. D.L. Shaw, & D. Weaver (Eds.), *Communication and Democracy* (pp. 137 – 151). Mahwah, NJ: Lawrence Erlbaum Associates Inc.
Wilson, D., Bingaman, J., Obozintsev, L., Brewer, P., & Paintsil, A. (2020). Media Messages and US Public Opinion about Artificial Intelligence.
Yang, J. Z., & Chu, H. (2018). Who is afraid of the Ebola outbreak? The influence of discrete emotions on risk perception. *Journal of Risk Research*, *21*(7), 834−853.
Zhang, B., & Dafoe, A. (2019). Artificial intelligence: American attitudes and trends. *Available at SSRN 3312874*.

Zhang, B., & Dafoe, A. (2020, February). US Public opinion on the governance of artificial intelligence. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* (pp. 187–193).

초록

본 연구는 컴퓨터 텍스트 분석 기술을 통해 인공지능 (AI)에 대한 뉴스 보도에 드러난 프레임과 감정을 분석하여 인공지능이 뉴스 미디어에서 어떻게 표현되는지를 살펴보는 것을 목적으로 한다. 프레임 추출을 위해 Structural Topic Model (STM) 기법을, 감정 추출을 위해 NRC Emotion Lexicon과 Linguistic Inquiry and Word Count (LIWC) 프로그램을 활용했다. 언론사의 정치 성향(보수 – 진보)과 미디어 유형(신문 – 방송)을 변수로 설정해, 추출된 결과와 각 변수와의 상관관계를 분석했다. 뉴스 미디어에 내재된 프레임과 감정을 파악함으로써, 그것이 AI에 대한 여론 및 태도 형성에 어떤 영향을 미치는지 예측할 수 있을 것이다.

주요어: 인공지능, 뉴스 미디어, 토픽 모델링, 데이터 분석 기법, 프레임, 감정

학번: 2019-25994