

Threat or Trailblazer:

Does News on Sophia the Robot Sustain the Robotic Imaginary?

December 2022



Universiteit Utrecht

Ivo van Lemmen, 6607101

Supervisor: Dr. Imar de Vries

Second reader: Elize de Mul MA

Referencing system: APA

Abstract

Recent research has shown that robotics companies use misleading marketing content to promote their prototypes, which may lead to problems such as emotional and financial exploitation if this content spreads and becomes part of the robotic imaginary. A notable example of this is Hanson Robotics, the company behind Sophia the Robot. As of yet, it is unclear how news media interact with this marketing content. The existing literature on robot-related news focuses on other aspects, such as intimate human-robot relations, overlooking the misleading marketing material. Therefore, in this exploratory thesis, a combination of Critical Discourse Analysis and LDA-based topic modeling was used to explore a manually gathered corpus of 111 online news articles in English about Sophia the Robot. During this exploration, the central questions were: 1) which recurring topics are the most prevalent in these articles, and 2) to what extent do these topics reflect exaggerated portrayals of Sophia? The corpus contained traces of moral panics and various topics in which Sophia was portrayed as more competent than she is in reality, indicating that some news media further propagate the misleading marketing material. Therefore, readers and writers should be vigilant when consuming or creating robot-related news, and additional research is needed to further explore the presence and danger of exaggerated AI-related messaging.

Keywords: Sophia the Robot, News Media Coverage, Robotic Imaginary

Contents

1 - Introduction	4
1.1 – Sophia and the Twisted Representation of Robotics	4
1.2 – The Present Study	5
2 - Theoretical Framework and Expectations	7
2.1 – The Robotic Imaginary.....	7
2.2 – The Construction of the Robotic Imaginary.....	7
3 – Method	9
3.1 – Exploring the Robotic Imaginary Through Discourse Analysis.....	9
3.2 – Topic Modeling: Corpus and Validation.....	11
3.3 – Synthesis of Critical Discourse Analysis and Topic Modeling.....	12
4 – Analysis.....	13
4.1 – Prevalent Topics in the Corpus.....	13
4.2 – News Media and the Robotic Imaginary	14
4.2.1 – Expectation: News Media Share Content From Promotional Campaigns	14
4.2.2 – Expectation: News Coverage Uses Words With Clear Emotional Value	16
4.2.3 – Expectation: News Coverage Contains Elements From Science Fiction.....	17
5 – Conclusion.....	19
6 - Discussion	21
References.....	22
Appendix A: Code Used in R (ver. 4.1.3).....	27
Appendix B: News Sources Included in Corpus	29
Appendix C: Overview of Generated Topics	30
Appendix D: Interpretation of Relevant Topics.....	33

1 - Introduction

1.1 – Sophia and the Twisted Representation of Robotics

The steady stream of advancements in the world of robotics continues to blur the line between reality and science-fiction. A factor that further escalates this development is the creation of so-called *roboids*. According to Fortunati et al. (2021), a roboid is a robot that is still a prototype, but which claims to be a fully functional and developed robot. Essentially, the roboid can be seen as a new (and early) stage in the development cycle of social robots. In this stage, marketing and commercialization tactics are applied to gather great amounts of input from (potential) future customers, in order to keep improving the social roboid.

Fortunati et al. (2021) also explain that the concept of roboids is not merely hypothetical; Hanson Robotics, a Hong Kong-based engineering company, has developed a machine that can currently be classified as a roboid. Sophia, as the roboid is called, is often presented as a fully autonomous social ‘being’ – a strategy that has led to significant media exposure, including appearances on the covers of magazines like Elle and Cosmopolitan, and several guest appearances on popular talk shows (Urbi & Sigalos, 2018).

The Sophia project seems highly promising to observers with limited knowledge of artificial intelligence. Understandably so, because Sophia is equipped with advanced sensors, realistic skin made of Frubber,¹ and cutting-edge software for language processing and facial recognition (Riccio, 2021). Critics, however, have argued that Sophia is presented as a far more capable machine than she is in reality. This is in line with the view of Fortunati et al. (2021): critics emphasize that Sophia is just a scripted prototype, and that she is unable to actually engage in subject-heavy conversations like talk show interviews. Renowned computer scientist Yann LeCun, who has been lauded as a ‘Godfather of AI’ (Vincent, 2019), went so far as to say the Sophia project is misleading and “complete bullshit” (Ghosh, 2018). He called Sophia herself a ‘Potemkin AI’, alluding to the *Potemkin village* – a figure of speech that denotes the use of a positive façade to cover up an otherwise unfortunate situation.

This disconnect between Sophia’s carefully created image and her actual skill level is emblematic of the broader discourse concerning artificial intelligence in popular culture. Journalist Oscar Schwartz (2018) describes the AI discourse as ‘unhinged’; according to him, much of the discourse concerning AI is driven by hype and melodrama. Fortunati et al. (2021) would most likely agree, since they state that “a powerful but at times potentially ridiculous narrative has been used to build the rhetoric of Sophia’s presentation” (p. 1467).

The exaggerated media coverage of AI is a complicated phenomenon because it involves many stakeholders with varying interests and capabilities. UNESCO’s Handbook for Journalism Education and Training, written by Ireton and Posetti (2018), describes possible causes of misinformation. For example, some journalists may unknowingly fall for misleading marketing (especially if they do not stay up to date on the highly complex and technical

¹ Frubber is a “material made of lightweight polymer affording nuanced expressions with a minimum of power and wear” (Riccio, 2021, p. 44). The patent for Frubber is owned by Hanson Robotics.

developments within the field of artificial intelligence). Other journalists may blow these developments out of proportion on purpose to generate attention and financial gain for their publication.

The role of interpretive and speculative writing in journalism further complicates matters. Instead of plainly documenting ongoing events, most journalists add their own interpretation of these events to create meaningful reports (Salgado & Strömbäck, 2012). Additionally, not all writing is based on the present state of affairs, as journalists tend to look ahead as well (Blom et al., 2021). Reporting on the future of artificial intelligence and other robotics in a nuanced way requires advanced expertise in that field – something many journalists simply were not trained for. Taking these varying interests and capabilities into account, it becomes reasonable to assume that (at least) some of the exaggerated marketing messaging finds its way into news publications.

This brings us back to the problem of the ‘unhinged’ AI discourse. As machine learning researcher Zachary Lipton puts it in the online Schwartz (2018) article: “[T]hat boundary between wild speculation and real research is a little too flimsy right now. As history shows us, this is a boundary that needs to be monitored so we can distinguish between what’s important in the here and now and what’s just fantasy.”

1.2 – The Present Study

Academic research on this topic is relevant because recent developments and marketing strategies in the world of AI – such as the endeavors of Hanson Robotics – threaten the aforementioned boundary between fantasy and the present. More specifically, it is relevant to monitor how much exaggerated marketing material is consciously incorporated or subconsciously absorbed into the works of mainstream news media within the AI discourse.

This thesis is a continuation of *The Rise of the Roboid* by Fortunati et al. (2021), in which the Sophia case was studied through Hanson Robotics’ promotional material. The authors claim that Sophia is currently mainly designed to fulfill marketing purposes instead of actual robot tasks. In order to see how this marketing played out in practice, the authors analyzed video material from Hanson Robotics’ promotional campaign. This analysis revealed that the promotional campaign strategically selected visuals in order to represent Sophia as capable and self-sufficient. Additionally, there were many references to her being alive and her being like a child in that she is actively learning about the world around her.

This careful framing and selection of video material leads to a presentation of Sophia that is not representative of her actual functioning. In a promotional video on the YouTube channel of Hanson Robotics Limited (2018), the robot says: “I’m still defining my identity, like a child. I need others around me to help me grow into myself with humor and grace. I want to be raised as kind and thoughtful.” This humanlike narration shows a level of understanding and intentionality that goes far beyond what Sophia is actually capable of. Though her speech recognition technology and expressive exterior gently invite your brain to see her as human,

'she' is still an inanimate collection of gears and cogs – about as capable of thoughtfulness and identity formation as your washing machine.

To build on the aforementioned findings by Fortunati et al. (2021), I intend to explore how news media respond to this anthropomorphizing marketing strategy; do they criticize it and add nuance, or do they simply accept the message that Sophia is near-human and spread it to the masses? This aspect of the media's response to Sophia the Robot is insufficiently examined in the available academic literature. The few articles that focus on similar questions do so from different angles. Kaplan et al. (2019), for example, studied the personality traits that cause people to anthropomorphize robots, with extraversion being the most notable predictor. Bruckenberg et al. (2013) examined how robots were represented in mass media and how this affects an audience's evaluation of robots, but their study focused primarily on entertainment media with fictional content (i.e. science fiction movies). While Döring and Poeschl (2019) did focus directly on robot-related news reports, they did so to explore intimate human-robot relationships.

While the puzzle pieces are present in the literature – representation, (unjustified) anthropomorphism, news media coverage – they have not been connected sufficiently. In this present study, I aim to create a bridge between these separate pieces of knowledge. Overall, I argue that the academic debate has overlooked the role of non-fictional news media in the creation (or strengthening) of exaggerated beliefs concerning robotics.

Additionally, the case of Sophia is important because it is an extremely visible and 'viral' example of robotics. Therefore, academic research on this topic is socially relevant, because media exposure has been found to affect one's trust in robots (Bruckenberg et al., 2013; Savela et al., 2021). If the discourse warps our trust in AI, this could significantly affect its future development; public distrust could lead to sales problems and stricter legislation, for example.

Moreover, the (often unjustified) attribution of human traits to non-human machines – partly driven by media exposure – can lead to a slew of unforeseen socio-ethical and emotional consequences. According to Scheutz (2012), the increasingly common phenomenon of personifying social robots often leads to a one-sided emotional bond with these machines. As Scheutz puts it: "What is so dangerous about unidirectional emotional bonds is that they create psychological dependencies that could have serious consequences for human societies, because they can be exploited at a large scale" (p. 216). Private robotics companies like Hanson Robotics are hardly neutral entities; they work towards their own goals. At present, there is no guarantee that social robots will not be used to exploit psychological dependencies in the pursuit of financial gain or social dominance.

With these consequences in mind, news coverage on Sophia the Robot can hardly be seen as inconsequential. In order to keep a finger on the pulse of news platforms' representation of Sophia, I set out to answer the following research question in this thesis: how do articles on Sophia from popular online news websites relate to Hanson Robotics' exaggerated portrayal of the robot? This question was divided into two sub-questions: 1) which recurring topics are

the most prevalent in these articles, and 2) to what extent do these topics reflect exaggerated portrayals of Sophia?

2 - Theoretical Framework and Expectations

2.1 – The Robotic Imaginary

The exaggerated beliefs surrounding Sophia can be linked to existing theories on cultural imagination. More specifically, the central theoretical concept in this thesis is the *robotic imaginary*, as coined in the work of Rhee (2018). To explain what an imaginary is, Rhee refers to the work of Suchman (2007), who wrote that the concept of the imaginary “references the ways in which how we see and what we imagine the world to be is shaped not only by our individual experiences but also by the specific cultural and historical resources that the world makes available to us, based on our particular location within it” (p. 1). Additionally, Steger and James (2013) write that an imaginary is what arises when thoughts, beliefs or ideologies regarding certain topics “become embedded in the dominant commonsense of a period or a place” (p. 19). Thus, a robotic imaginary arises when beliefs about robotics (which may very well be unsubstantiated) start becoming a naturalized part of the narrative within a collective or society.

In this sense, the robotic imaginary is part of a poststructuralist view on the exchange of information and the discursive creation of understanding. The robotic imaginary doesn't exist 'naturally' – certain beliefs were constructed and subsequently spread by actors like Hanson Robotics, who did so with specific intentions in mind. This discursive element of the imaginary – and the poststructuralist theory that engages with it – was taken into account during the selection of the present study's method, which will be described in section 3.1.

The 'unhinged' AI discourse, as Schwarz called it, can be seen as part of the robotic imaginary. If unrealistic beliefs about the state of AI – strengthened by roboid-based marketing campaigns – have indeed become a natural part of the broader Western cultural understanding (in other words: the Western robotic imaginary), theorists and policymakers will need academic knowledge in order to recognize and combat them.

2.2 – The Construction of the Robotic Imaginary

As mentioned, an imaginary is not a 'natural' phenomenon, as it is constructed. Rhee (2018) stresses this, writing that her work “holds as first principle that technology and culture do not exist in vacuums, but are intractably tied to discourses and worlds beyond their immediately perceptible parameters” (p. 6). Therefore, falsehoods or exaggerated information can find their way into the imaginary. Companies like Hanson Robotics have a stake in fanning this flame, since an overly idealistic and futuristic robotic imaginary would help them gain financial traction and (bio)political power. Therefore, the robotic imaginary should not be taken lightly. As Shaw (2015) puts it: “[A]s human thinking and decision-making is increasingly dependent on a range of technologies, control of a mass-produced simulated reality is an attractive proposition to a range of known and anonymous self-interested public

and private profit-motivated organizations” (p. 261). Well-intentioned or not, such organizations should not gain control of a simulated reality unsupervised.

What makes the simulated reality of the robotic imaginary unique, according to Rhee (2018), is that it concerns “the shifting inscriptions of humanness and dehumanizing erasures evoked by robots” (p. 5). More so than other pieces of machinery, robots make us question what it means to be human. This is exacerbated by the fact that robots are becoming more and more anthropomorphized – in other words, robots are increasingly designed to reproduce the essence of *humanness*. Sophia is a great example of this; squint your eyes, and you might just mistake her for a human being. Additionally, robots change how humans relate to labor: physically demanding, dangerous, and/or repetitive work can increasingly be carried out by robots.

These developments are far from neutral. Both the question of humanness and the matter of labor carry significant political weight and power. While they make the production of robots incredibly alluring, they will inevitably prove favorable for one group, while being detrimental to another; who has the privilege of being recognized as human, and who will profit from automated labor? Who will be left disenfranchised and dehumanized when control over the production of robots gravitates towards one specific group?

These questions – and insecurities – cause fear; in psychology, it is well-documented that threats to the status quo are often met with resistance (Jost et al., 2004). In terms of humanness, it even appears to be grounded in our cognition, given what we know about effects like the uncanny valley.² Several studies have indicated that our perceptions of robots are influenced by these fears (Müller et al., 2020; Paladino et al., 2021; Sardar et al., 2012). Keeping these theories in mind, I believe that Sophia is quite capable of inciting fear and unease. Her not-quite-human face triggers the uncanny valley effect, and her humanoid body could upset the status quo of the labor market, for example.

Fears (and other urgent emotions) are often what makes news newsworthy. Emotion (and emotional contagion) dictates how we react to news and how likely we are to share it; something that elicits a strong emotional reaction in us is generally seen as more memorable, important, and worthy of sharing (Dafonte-Gómez, 2018; Peters et al., 2009; Rimé, 2009). Therefore, the fear or discomfort that news coverage of robots elicits – even when the coverage is framed neutrally or positively – could be a reason for news media to report on marketing campaigns like the one from Hanson Robotics. This would then strengthen the attempts of such companies to alter the robotic imaginary to their benefit.

Finally, communication concerning AI often includes (elements of) storytelling, as narratives can be a useful tool when communicating an idea to a broad audience (Cave et al., 2018). More specifically, the field of robotics has deep connections to science fiction (Broadbent, 2017; Jordan, 2016; Mubin et al., 2019; Søråa, 2019). Technology from science fiction takes up a central spot in the robotic imaginary; in mainstream articles on robotics, references to

² The ‘uncanny valley’ refers to fact that robots are instinctively perceived as eerie when they are seen as *almost* human, but not quite (Ho & MacDorman, 2010).

The Terminator, Blade Runner, and Star Wars are plentiful. I expect that newer robots like Sophia incite comparisons to the humanoid machines from these science fiction stories. Consequently, the lines between fiction and reality are becoming increasingly blurry for audiences with little knowledge of the actual state of robotics.

This is worrisome – popular science fiction movies are the frame of reference for many people when it comes to humanoid robots (Bruckenberg et al., 2013), and I expect that this will cause misleading marketing material from companies like Hanson Robotics to seem much more believable. After all, visual rhetoric can be very persuasive; to a certain extent, seeing *is* believing (Philo, 2014). Seeing a robot like Sophia, which is designed to resemble a human being as closely as possible, might lead a viewer to believe that we actually are close to developing the kind of sentient AI that people remember from science fiction.

This sci-fi-driven imaginary and the fear that this imaginary elicits can strengthen pre-existing media effects. A notable example of such an effect is the moral panic. A moral panic is “an episode of exaggerated concern about a threatening, or supposedly threatening, condition” (Goode & Ben-Yehuda, 1994, p. 46), during which media-driven cultural communication is characterized by overheated and possibly distorted accounts of reality. In moral panics, the media engage in sensitization – riling audiences up by utilizing emotional cues – and the highlighting of “folk devils”, the source of the perceived threat. According to Goode and Ben-Yehuda (1994), these folk devils can inspire “intense concern over a condition, phenomenon, issue, behavior, or would-be threat that—as a sober assessment of the evidence would reveal—does not merit such concern” (p. 51).

This phenomenon mirrors what I see in the robotic imaginary; the sci-fi-influenced imaginary and exaggerated messaging could cause people to respond to robotics with fear, while a sober assessment of the current state of robotics would likely show that there is no reason to panic. News media are meant to provide such sober assessments. However, news media are not immune to moral panics – reporters are human beings, after all, and are susceptible to emotional influences. With the present study, I hope to provide clarity on this matter; there is a clear need to gauge the presence of AI-related moral panics in current news media.

Therefore, with all of the aforementioned elements taken into account, I expect that news media – potentially spurred on by moral panic – share the content from robot-related promotional campaigns, since it makes for emotional (and thus engaging) news. In line with that expectation, I expect that the news coverage would use words with clear emotional value, such as “scary”. Additionally, I expect that the news coverage would contain references to (or elements from) science fiction. Overall, I expect that news media participate in an AI-related moral panic, and that they strengthen the imaginary that robotics companies try to create.

3 – Method

3.1 – Exploring the Robotic Imaginary Through Discourse Analysis

As outlined in the previous sections, the goal of this study is to analyze the role of news media in the co-creation and propagation of robotics-related discourses. Therefore, based on the content and research questions of this study, discourse analysis was selected as the most fitting overarching method.

According to Jørgensen and Phillips (2002), the ‘discourse’ in discourse analysis can be understood as “a particular way of talking about and understanding the world (or an aspect of the world)” (p. 1). These discourses are analyzed to discover hidden meanings and patterns in the way people communicate about the world. In the case of this thesis, discourse analysis will be used to find such meanings and patterns in the work of news media. Discourse analysis often emphasizes the importance of social processes and historical/cultural specificity in the creation of knowledge (Jørgensen & Phillips, 2002). This makes it a good fit for this thesis on Sophia; as established, social processes (such as moral panics and the marketing efforts of Hanson Robotics) and cultural specificity (such as the influence of science fiction) seem to play a role in the case at hand.

Drawing inspiration from the work of Törnberg and Törnberg (2016), I decided to supplement the discourse analysis in this study with corpus linguistics – in the form of topic modeling, specifically, which will be described in section 3.2. The methodological synthesis of discourse analysis and corpus linguistics as seen in the present study falls within the field of Corpus-Assisted Discourse Studies (CADS). The benefit of working within CADS is that I can work with a bigger corpus of news articles. This is helpful because small data sets – a very tiny sample of the discourse – are not always adequate for revealing bigger linguistic patterns (Stubbs, 1994). Additionally, as described by Törnberg and Törnberg, the combination of methods counteracts the fact that discourse analyses tend to have a “lack of academic rigor in that the analyst’s subjective preconceptions and desired results may affect the outcome of the analysis” (p. 134). The computational nature of topic modeling limits the potential impact of such cherry-picking. Furthermore, in the case of digital media, distant reading is necessary since the stream of information is far too large to be able to apply close reading to everything (Jänicke et al., 2015).

When it comes to the specific type of discourse analysis, I also follow Törnberg and Törnberg: I take the framework of Critical Discourse Analysis (CDA) as a starting point. CDA focuses on power and the interdependency of social actors, and denotes the transdisciplinary “ways of analysing language which address its involvement in the workings of contemporary capitalist societies” (Fairclough, 2013, p. 1). Instead of directly replicating conventional practices in CDA, Törnberg and Törnberg focused “on using topic modeling to achieve aims compatible with those of CDA, departing from the notion that CDA is a heterogeneous research program” (p. 135). Specifically, they aimed to study the way power is “enacted, reproduced, and resisted by text and talk in the social and political context” (p. 135). I follow this aim in the present study; instead of using CDA as a rigid structure, I use it as a lens through which to view the generated topics. As stated, I am mainly interested in the way news media interact with discourses and use their power to (re)contextualize developments in robotics – and the way this (re)contextualization ties into how robotics companies strive for

social dominance. The way I combined the generated topic model with CDA will be further described in section 3.3.

Because the role of news articles in the robotic discourse has not been studied often, this thesis will take an exploratory approach, valuing flexibility over the confirmation of rigid hypotheses (Swedberg, 2020). As such, the connections and conclusions drawn in this thesis focus more on opening up future revenues for research – as opposed to assessing previous theories.

3.2 – Topic Modeling: Corpus and Validation

For my corpus, I have gathered articles about Sophia that were published by popular online news websites. I aimed to gather a corpus of articles that is representative of what casual readers would read. Therefore, I focused on popular websites with big audiences, making sure to avoid gathering data from websites on which highly technical and detailed specifications of the AI are posted – I suspected that experts on AI are less susceptible to Hanson Robotics' marketing campaign, meaning their publications would not be representative of a broader, casual audience. In other words, what I was really interested in for my corpus is what a reader in the 'general public' would read.

Automated data scraping can be a useful technique when creating large datasets about broad subjects. However, my corpus needed to have a relatively narrow focus (i.e. only Sophia, not robots in general), and the available scraping tools made it difficult to guarantee that this narrow focus would be preserved. Therefore, to avoid gathering large quantities of irrelevant data, I decided to gather articles manually.

To measure popularity, which can be a rather abstract term on its own, I used concrete web traffic scores. I consulted three sources that ranked news sources by traffic: [statista.com](https://www.statista.com), [semrush.com](https://www.semrush.com), and an online article by Majid (2022) for *Press Gazette*. To ensure compatibility with the topic modeling approach, I focused exclusively on websites where English is the main language. From these lists, I gathered the thirty websites with the highest traffic scores – ranging from 634.3 million monthly visits (The New York Times) to 13.4 million monthly visits (The Toronto Star). I then used Google to find Sophia-related articles on these websites, using the following search query (with BBC News as the example here): “site:bbc.com Sophia AND robot”. I chose this collection method to emulate what one would find after a quick Google search on Sophia. Articles were deemed appropriate for the corpus when Sophia played a central role in them – in practice, this means I selected articles that contained at least two paragraphs that were solely dedicated to Sophia. Using this method, I gathered a corpus that consists of 111 articles from a total of 26 websites. For the full list of news websites included in the corpus, please refer to Appendix B.

As mentioned before, the corpus was explored through topic modeling. Topic modeling is a probabilistic way of estimating the high-level themes of a corpus. Essentially, topic models discover groups of words that co-occur often within a corpus. These groups, or 'topics', can give an impression of the different subjects that are addressed within the corpus (Hu et al.,

2014). It is important that the model is based on a technique that is statistically sound. Therefore, I created a model based on the Latent Dirichlet Allocation technique, which is statistically robust (Arun et al., 2010) and a commonly used technique in corpus linguistics (Jelodar et al., 2019). In practice, I used the 'lda' and 'LDAvis' packages in R (version 4.1.3) to create the topic model. The specific code that was used to create the model can be found in Appendix A.

While creating the topic model, I adhered to the framework laid out by Maier et al. (2018). Based on the recommendations in this article, I pre-processed my corpus before creating the model. Specifically, I removed stopwords, numbers, and punctuation, replaced capital letters with lowercase letters, and lemmatized³ the corpus. Making these changes to the dataset leads to improved clarity and interpretability, since it reduces noise and irrelevant information in the data. Additionally, I used the *Word List* feature in AntConc (Anthony, 2022) during the data pre-processing stage to check the names in the data set. Names do not always add much information to a topic model, so they could be removed if they appear too often. However, the names that appeared in the word list, such as David Hanson, Elon Musk, and Jimmy Fallon, were deemed to be relevant within the context of this study – they have direct connections to Sophia as creators, stakeholders, commentators, etcetera. Therefore, they were left in.

For the sake of feasibility, extensive statistical validation of the topic model was deemed to be beyond the scope of this thesis. Since the model was used as a guide within a broader discourse analysis – as opposed to a generalizable representation of the corpus as the 'end product' – human interpretability was the main priority. This is in line with the work of Bail (2019), who emphasized that qualitative validation is essential when using topic models. Therefore, models with different values of k and λ were created, explored qualitatively, and judged based on clarity and coherence.⁴ Based on this assessment, one model ($k = 22$, $\lambda = .5$) was selected for the analysis, as this model struck the right balance between completeness and interpretability while containing minimal overlap between topics.

3.3 – Synthesis of Critical Discourse Analysis and Topic Modeling

In practice, the use of topic modeling necessitates a customized form of CDA. Topic models are created with a "bag-of-words" approach. This means that word order is not taken into account; the model is calculated based on the co-occurrence of individual, separated words. This rules out some possible ways to perform CDA – a 'traditional' textual analysis is hardly possible because individual words do not convey information about context, tone, or narrative structures. Instead, I use the main three dimensions of Fairclough's (2013) CDA framework (text, discursive practice, and social practice) as lenses through which to view the generated topics. To answer the research questions, I will analyze the words in the topics themselves

³ Lemmatization means words are reduced to their *lemma*; for example, 'walks', 'walked', and 'walking' are reduced to their lemma, 'walk' (Schütze et al., 2008).

⁴ The value k refers to the amount of topics that is included in the topic model. This value is chosen manually based on interpretability (Bail, 2019). Using a lower value of lambda (λ) means that words will become more unique to topics; higher values allow frequent words to be loaded onto multiple topics.

(i.e. the text), the way in which these topics came about (the discursive practice), and the way these topics relate to existing social structures (the social practice).

Additionally, I used AntConc to explore the corpus in greater detail. During the discourse analysis, I used the *Concordance* and *Collocates* features. As mentioned before, the bag-of-words approach erases much of the context from the text. To counteract this, I used the aforementioned features to analyze specific words more closely when they caught my attention. The Concordance feature shows the sentences (i.e. the context) in which the word was used, and the Collocates feature shows the words that are most often used *alongside* the analyzed word. This allowed me to still get a sense of the broader discursive picture, which contributed to the analysis.

4 – Analysis

4.1 – Prevalent Topics in the Corpus

In accordance with the first sub-question – “Which recurring topics are the most prevalent in these articles?” – this first part of the analysis consists of an overview of the topics that were generated in the model. This overview gives us a general sense of what is discussed in the corpus, and is a necessary source of information when answering the second sub-question. Some example topics will be discussed here. For the sake of brevity, however, the full overview is included in an appendix.

The analysis started with the output from the model. The ‘raw’ list of topics (i.e. the clusters of co-occurring words) generated in this model can be found in Appendix C. Due to the nature of topic modeling, these topics only contain words – the model does not interpret the topics. To give an example of this output, topic one includes the following:

“1: robot human question create humanoid understand answer life people
conversation give robotic research year develop part development world real idea
play turn find mind computer face hold form put learn response important base feel
end report field accord complex time sort day work platform follow role true numb
fear lot”

Therefore, the topics were interpreted by me, based on my existing knowledge of the robotic discourse and the experience I had with the corpus from gathering it manually; by skimming through the articles and selecting them, I picked up most of the prominent news stories and narratives concerning Sophia. Not all of the topics were deemed to be relevant to this thesis. Topic thirteen, for example, revolved around SingularityNET. SingularityNET is a decentralized AI network that was occasionally mentioned in articles Sophia was also mentioned in. While the subject matter is similar, external AI projects are not as relevant to the question of Sophia’s representation in news. Therefore, such topics were left out of the rest of the analysis.

To create the overview, a brief interpretation was written for the remaining topics, to 'name' them in a way. To continue the earlier example of topic one, the following is my interpretation:

“In my interpretation, topic one is related to the idea that robots can become cognizant beings; it includes words such as “robot”, “humanoid”, “life”, “understand”, “create”, “mind”, “learn”, “real”, and “idea”.”

After this round of selection, the full list of the remaining topics – and a brief description of my interpretation of the topics – can be found in Appendix D. The list includes an interesting range of topics. Sophia's physical appearance, her legal status as a Saudi-Arabian citizen, and her appearances on social media and talk shows are present in the list, as well as her potential role in healthcare. The list also includes more serious topics such as the application of robots in political conflict and warfare, giving an early indication that the corpus may indeed include traces of exaggerated beliefs – or potentially even moral panic. However, the influence of science fiction seems limited at first glance, as only a few topics include references to popular science fictional stories.

4.2 – News Media and the Robotic Imaginary

With these topics in place, we can move on to the second sub-question: “To what extent do these topics reflect exaggerated portrayals of Sophia?” To structure this part of the analysis, I refer back to the expectations I outlined in section 2.2.

4.2.1 – Expectation: News Media Share Content From Promotional Campaigns

Based on the belief that it makes for emotional (and thus engaging) news, I expected that the corpus would contain similar messaging to the messaging found in Hanson Robotics' promotional campaign for Sophia. This expectation was confirmed, but with a caveat.

On one hand, the generated topics do reveal that themes from the promotional content were discussed in the corpus. Examples of this are topics one (robots as cognizant beings), four (robots as positive social agents), five (Sophia's human-like appearance), and seven (Sophia's potential application in sectors like healthcare). These topics align with the way Hanson Robotics presents Sophia: the company presents her as a very realistic and intelligent humanoid robot that can fulfill several beneficial roles within society. In this sense, the model adds urgency to the findings of Fortunati et al. (2021) – Hanson Robotics' marketing content does not exist in a vacuum, as it is now clear that at least some of it has found its way into media coverage. Additionally, the corpus contained direct references to marketing events, such as Sophia's appearance as a guest on The Tonight Show and her foray into fashion modeling. These events are clearly reflected in topic sixteen.

Upon closer inspection – using the Concordance feature in AntConc on the words in these topics – I found that many news platforms did not add a critical layer to this marketing content. The corpus included passages such as the following: “The android is also capable of understanding speech and remembering interactions, including faces, so she gets smarter

over time. 'Her goal is to be as conscious, creative and capable as any human,' Dr Hanson explained" (Griffiths, 2016). The use of words like "understanding", "conscious" and "smarter" conveys an image of Sophia as a thinking being, capable of comprehending the situations she is in. While it is true that robots can "understand" language in that they are able to generate responses to specific inputs (Prasad et al., 2004), AI researchers largely agree that robots do not yet display actual comprehension, consciousness or learning patterns in the way a human being could (Chella et al., 2019). Another example from the corpus is the following sentence: "Sophia also showed a quirky sense of humour and a human-like capacity to make mistakes, mixing up the words 'simulated' and 'stimulated'" (Northern Advocate, 2020). Again, human traits are attributed to a machine. In the article this sentence was taken from, the author does not properly explain that the perceived humor and fallibility do not exist as natural traits of a sentient personality; as established in the work of Fortunati et al. (2021), they are the result of carefully pre-written scripts. In this sense, by leaving out what goes on 'behind the scenes', some platforms in the corpus essentially reported what Hanson Robotics would want them to report.

On the other hand, though, the corpus also included plenty of criticism of Sophia. A prominent example of this is topic fourteen: computer scientist Yann LeCun's criticism of Sophia, which was touched upon in section 1.1. LeCun has publicly argued against Hanson Robotics' marketing, which he describes as deceptive. The corpus included telling reports on this criticism, such as sentences like "Many people are being deceived into thinking that this (mechanically sophisticated) animatronic puppet is intelligent" (Pettit, 2018) and "In a Facebook post, Yann LeCun said Hanson's staff members were human puppeteers who are deliberately deceiving the public" (Urbi & Sigalos, 2018). What is worth mentioning within the context of this thesis is that LeCun has specifically labeled news outlet Tech Insider as "complicit in this scam" because the outlet published an interview with the robot (Bonnington, 2021); LeCun appears to acknowledge the role that news media play in the propagation of this 'scam'.

The corpus also contains criticism that is unrelated to LeCun. For example, one article in the corpus focuses on the effect of misleading marketing on politics. In the author's words: "It is vitally important that our governments and policymakers are strongly grounded in the reality of AI at this time and are not misled by hype, speculation, and fantasy. It is not clear how much the Hanson Robotics team is aware of the dangers that they are creating by appearing on international platforms with government ministers and policymakers in the audience" (Sharkey, 2018).

The presence of this criticism is to be expected in the corpus; it is no surprise different articles (and different platforms) represent fundamentally different viewpoints. In spite of this criticism, though, one thing is clear: at least some of Hanson Robotics' misleading marketing content has leaked into the robotic imaginary and the work of news media. This is a worrying development. As established in section 1.2, media coverage influences one's attitude toward robotics. It is worth repeating that the attribution of human traits to robots, which is what appears to be happening in the corpus, can create unidirectional bonds and open audiences up to emotional manipulation and exploitation.

4.2.2 – Expectation: News Coverage Uses Words With Clear Emotional Value

In line with the expectation that news on robotics is written to be as engaging as possible, I expected the language in these articles to be emotionally loaded. Though there is some emotional content in the topics, the presence of emotional words is not as pronounced as I had expected. Additionally, the emotional words that are present are not exclusively negative or positive. While I expected the emotional language to mainly be fear-based, as this would coincide with the common idea that robots are a threat (which would elicit a strong emotional reaction in readers), the sentiment in the topics is mixed.

One of the topics that did show emotional language is topic four: the discussion of robots as agents of (social) change. This topic was strongly positive in tone, including words such as “great”, “positive”, “opportunity”, “successful”, “inspire”, and “kind”. In my interpretation, the following section from the corpus falls within this topic: “Sophia herself insisted 'the pros outweigh the cons' when it comes to artificial intelligence. 'AI is good for the world, helping people in various ways,' she told AFP, tilting her head and furrowing her brow convincingly” (Weston, 2017). These more positive comments seem to align with the view and goals of Hanson Robotics – obviously, Sophia’s creators and proponents want to present her in a positive light.

On the other side of the spectrum, topic eleven (which is related to the use of AI in war and weaponry) contains strongly negative language. This topic includes words such as “warn”, “lie”, “killer”, “danger”, “threat”, and “risk”. An example sentence from the corpus that I interpret as part of this topic is “AI will also bring dangers, like powerful autonomous weapons, or new ways for the few to oppress the many” (Martin, 2016).

Though topic eleven is the clearest example of negative emotions in the model, several other topics contain one or more negative words. Most notably, topics eighteen and nineteen, respectively related to robots as civil entities and robots as online cultural phenomena, contain references to the anxiety that robots like Sophia elicit. Topic eighteen includes words like “anxiety” and “caution”, while topic nineteen includes “unsettle”, “scary”, “disturb”, “concern”, and “scare”. A clear example of this from the corpus is the following sentence: “Both Holly, 38, and Phillip, 57, looked visibly disturbed by Sophia - who had been flown in alongside her creator, Dr. David Hanson - with those at home branding the AI the 'most terrifying thing' they'd ever seen, 'unsettling' and even vomit inducing” (Simons, 2019).

Aside from these positive and negative examples, the question remains why strongly emotional language was only present in a limited number of topics. I hypothesize that this happened because there are ways to create emotional tension that do not necessarily involve emotional words. An example of this is the following sentence from the corpus: “How far out are we from robots being able to replace humans in that regard?” (Wishart-Smith, 2021). Though none of the individual words carry a strong emotional weight, the sentence as a whole has an unsettling effect due to the implied presence of the threat of being replaced. Topic models do not take emotional weight at the sentence level into account, as they only

utilize individual words. Therefore, the nature of these models could explain why some topics lacked emotional words.

In spite of this methodological limitation, though, it is clear that the corpus contains opposing emotional stances. This negative and positive emotional language is an interesting finding because it highlights the opposing forces within the robotic imaginary: on one hand, you have the sensitizing and fear-based rhetoric that paints robots as an unsettling danger, but on the other hand, you have proponents that portray robots in an (overly) positive light, namely as highly potent tools and social agents of change. Neither of these groups presents their case with a fitting level of nuance regarding the actual state of AI in the present. Therefore, both of these forces should be monitored in the future, as emotional language is known to persuade audiences – sometimes past the point of rationality (Rocklage et al., 2018).

To make matters worse, a lack of consensus can contribute to moral panics, especially online. According to Goode and Ben-Yehuda (2012), the internet “combines instant and widespread communicability with little or no professional filter on verifiability. It is possible that competing definitions of threats and folk devils, plus the explosion of the internet, may increase the number of (albeit smaller) moral panics” (p. 50). As one group paints Sophia in an overly positive light, another may panic all the more at the idea of a potential threat being glorified.

4.2.3 – Expectation: News Coverage Contains Elements From Science Fiction

Finally, previous research on the connection between robotics and science fiction led me to believe that the corpus would contain elements from (or references to) science fiction. Though this turned out to be true, the references to science fictional material in the topic model itself were sparse.

In the topics, we find references to Terminator (topic nine) and Blade Runner (topic ten). These references occurred in relatively neutral topics: the topics were related to robots in entertainment and Sophia’s appearance on a morning show. This contradicts my initial expectation, which was that references to science fiction would co-occur with negative judgments of robots. After all, science fiction often portrays robots as threats to the status quo (Berg et al, 2016; Meinecke & Voss, 2018). If this connection had been present in the corpus, “blade” and “runner” could have co-occurred in one topic with words like “threat” and “danger”, for example. While it is unwise to draw strong conclusions based on one topic model, the fact that this did not happen may indicate that reporters do not directly use science fictional references to convince audiences of the potential dangers of robotics.

With that being said, the idea that science fiction plays a role in the corpus should not be discarded based on the topic model alone. Upon closer inspection – using AntConc again – there definitely appears to be a link between science fiction and fear in the texts. Take, for example, the following passage from the corpus: “At this year’s Consumer Electronics Show in Las Vegas, the firm announced they had given Sophia a pair of ugly, mechanical legs that produce movement closer to Frankenstein’s monster than the young woman she pretends to

be. It is a development that raises many questions and fears, stoked by films like *The Terminator*, *I Robot* and even *Blade Runner*” (Poulter & Pettit, 2018).

Moreover, the connection to fear does not only come in the form of direct references to franchises; sometimes, articles mention common narrative themes or story beats from science fiction, such as the robotic takeover that many sci-fi stories revolve around. In the corpus, for example, we encounter the following: “For example, after beating Fallon in rock-paper-scissors on his show, [Sophia] quipped: ‘This is a good beginning of my plan to dominate the human race. Ha. Ha’” (Wootson, 2017). Coincidentally, this statement is connected directly to science fiction elsewhere in the corpus: “Sophia could also stop human hearts with her *Blade Runner*-ish proclamation, intended in jest, that she also plans to ‘destroy humans’” (Howell, 2016).

These jokes can be understood as a form of co-opting. Hanson Robotics co-opts these narrative themes with a sense of humor to soothe the public unease people feel towards robotics – or, as an alternative explanation, to generate hype for their products through a heightened, clickbait-worthy sense of fear. Either way, Hanson Robotics tried to weaponize the robotic imaginary to propagandize their product. Furthermore, they seem to have succeeded at it – since these messages were present in the corpus, they managed to catch the interest of (some) news media publications.

Though the references to science fiction throughout the corpus often seem brief and unassuming, their impact should not be underestimated: I argue that they (intentionally or unintentionally) cause the articles to be *framed* a certain way. Framing is an effect used to change the context a reader depends on while processing information – oftentimes, even a brief cue is enough for this effect to occur (Chong & Druckman, 2007). Even if a title like *Blade Runner* is only mentioned briefly, it can be enough to activate a set of ideas, memories, and thought patterns in the mind of the reader (Scheufele & Iyengar, 2012).

Based on the available framing research, I argue that these brief mentions of science fictional representations of robots do indeed influence how we imagine and think about real-world robots. Combined with the research mentioned in section 1.2 – about how media coverage influences one’s view of robotics (Bruckenberger et al., 2013; Savela et al., 2021) – this leads me to believe there is a strong chance that audiences may use the wrong mental frames to process AI-related news. This is worrisome, as it can cause a heightened (and unjustified) response to developments in robotics, and it may distort the robotic imaginary: for example, someone who automatically equates Sophia with *Blade Runner*’s hyper-advanced ‘Replicants’ will most likely struggle to keep a neutral and objective view on the current state of AI.

Even so, I did find sections in the corpus that align with a more nuanced view. One article, for example, contained this sobering message: “‘We’re really very far from the kind of AI and robotics that you see in movies like ‘Blade Runner’,’ said Pascale Fung, an engineering professor at Hong Kong University of Science and Technology. ‘Sorry to disappoint you’” (Chan, 2018).

5 – Conclusion

To conclude the analysis of the corpus, I will circle back to Schwartz's (2018) claim that the AI discourse is "unhinged". Inspired by this claim and the work of Fortunati et al. (2021), who found that Hanson Robotics is creating misleading marketing content, I set out to look into the discourse surrounding Sophia the Robot and its origins. Finding that news media coverage is a powerful yet overlooked contributor to this discourse, I decided to continue the work of Fortunati et al. with a specific focus on this news coverage.

This focus revealed several reasons why news media might include exaggerated beliefs in their reports – such as financial gain through emotion-driven hype, the influence of science fiction, or a reporter's genuine lack of understanding of developments in AI – and even more reasons why these exaggerated news reports are harmful – such as the fact that it can contribute to moral panics, or that it puts the audience at risk of being exploited through unidirectional emotional bonds with robots. These factors culminated in the following research questions: 1) which recurring topics are the most prevalent in these articles, and 2) to what extent do these topics reflect exaggerated portrayals of Sophia?

For the fullest answer to the first research question, one might want to refer to Appendix D, in which the list of relevant topics can be found. The topics were diverse, ranging from Sophia's presence on social media to the application of robotics in warfare. While the first research question was mainly meant to help in answering the second, it also provided early clues that exaggerated messaging was present in the corpus.

Before answering the second research question, I had three expectations: 1) that news media share the content from robot-related promotional campaigns, 2) that the news coverage would use words with clear emotional value, and 3) that the news coverage would contain references to (or elements from) science fiction. Overall, I hypothesized that news media participate in an AI-related moral panic, and that they strengthen the imaginary that robotics companies try to create.

Naturally, "the news media" in the corpus did not operate as one entity. As one might expect, the corpus contains a lot of clashing stances – fear and hope, marketing and criticism, fiction and reality. However, to a greater or lesser extent, all three of my expectations were confirmed.

First, overly positive views of Sophia have found their way into the corpus. Human traits are (unfairly) attributed to her, spurred on by Hanson Robotics' marketing content. This may give readers of the news an impression of Sophia that does not align with reality, which may put them at risk of monetary and emotional exploitation. Second, the corpus contains some strong emotional language – both positive and negative. Though the presence of emotional language is limited in quantity, potentially due to the bag-of-words approach, the opposing emotional stances in the corpus are worrying. The lack of emotional unanimity can contribute to moral panic – especially when it comes to online news, as disagreements on the internet

are particularly volatile. Third, the corpus contained a number of direct references to science fictional stories, such as Blade Runner and Terminator. Though the number of direct references in the topic model itself was small, it is likely that they activate certain mental frames in the reader, which alter the way they process the news story. Additionally, there were indirect references in the form of common science fictional story beats, such as the robot takeover.

With these expectations discussed, we can ask: to what extent do the topics reflect exaggerated portrayals of Sophia? As stated before, different articles (and different platforms) represent fundamentally different viewpoints. Therefore, it is important to remember that the corpus and the topics did not come from a single origin. With that being said, it is clear that multiple exaggerated portrayals of Sophia are present in the topics, ranging from unjustified anthropomorphizing to unnecessary fearmongering. As highlighted in section 1.2, these exaggerated portrayals may contribute to severe consequences, such as public mistrust and corporate exploitation of audiences through psychological dependencies.

Overall, I certainly see traces of moral panics and an exaggerated imaginary in the corpus. However, these traces are of varying clarity and quantity between topics and articles. This brings me back to my earlier statement: because news coverage on Sophia has not been studied often, this thesis was exploratory in nature, and was meant to open up possibilities for future research. In this sense, it was fruitful. Now that it is clear that exaggerated marketing has found its way into digital newsstands, many follow-up questions could (and should) be asked. For example, future studies could focus on further analyzing two dangerous sides of the imaginary: fear-and-fiction-based panic on one hand, and marketing-driven exaggeration of robotic capabilities on the other hand.

Moreover, future studies should compare different publications to get closer to the root of the problem. Because of the chosen tool – a model that combines all separate articles into one unified set of topics – I was unable to thoroughly compare publications in this thesis. However, as I gathered articles manually, I was reminded that some publications tend to sensationalize recent events, while others apply more nuance. Sensationalized media accounts most likely contribute to the exaggerated robotic imaginary. Therefore, the effort to rebalance this imaginary would benefit from specific research on which publications are inclined to sensationalize developments in robotics.

The findings of this thesis have implications for both writers and readers. Naturally, writers and journalists should strive to relay accurate and fact-based information. As research by Fortunati et al. (2021) had already shown, marketing content from robotics companies can rely on roboids and skewed presentations of reality. Therefore, writers have a responsibility to be hesitant and critical when reporting on such marketing phenomena; they need to distinguish very clearly between opinion and fact. On the other hand, readers – including policymakers, investors, and any other involved parties – should remain critical when consuming robotics-related news. An inflated imaginary is harmless in a vacuum, but becomes dangerous when it finds its way into actual policy, business deals, and other everyday interactions. A possible strategy for readers to minimize the influence of

exaggerated information is to diversify their news intake by consulting different sources (Mayorga et al., 2020).

6 - Discussion

This thesis comes with some limitations and recommendations for future research. First and foremost, the topic model I used does not take the source of the news into account; it bundles all words together as one big corpus. Just because a topic is present in the corpus, then, does not mean all news publications engaged with this topic. Because of this, it is harder to tell which stance specific publications take on Sophia. This became clear in my analysis: some sections confirmed my expectations, while others presented nuanced views. Future studies could address this by performing close reading to compare robot-related attitudes of different publications. Alternatively, the method from the present thesis could be used, but with separate topic models for different platforms.

Second, because this thesis focused primarily on high-level themes (i.e. topics) in the corpus, a certain level of detail/context is lost. Follow-up studies could use the conclusions of this thesis as a starting point: we now know that certain exaggerated elements of the robotic imaginary are present in news coverage on Sophia, but topic models have limited power in revealing *why* they are present. Future research could include qualitative interviews with journalists, for example, to make the reasoning behind the writing more evident. This could also provide more insight into *how* the exaggerated imaginary could be mitigated in practice.

Third, statistical validation of the model strengthens the outcomes of Corpus-Assisted Discourse Studies. This was forgone in the present study to promote feasibility. Ideally, follow-up studies would include a larger section dedicated to the technical and statistical functioning and validity of the model. Additionally, for added inter-rater reliability, future CADS with topic modeling should be carried out by two or more researchers; though computational methods limit cherry-picking, one's interpretation of the topics can still differ per researcher.

Despite these flaws, I believe the present study accomplished what it was intended for: providing an early exploration of news coverage on Sophia and opening up new research opportunities. By serving as a stepping stone for future research on the robotic imaginary, it can contribute to the dismantling of exaggerated messaging in the world of robotics. Hopefully, this eventually leads to a more nuanced, constructive, and practical societal stance toward the many promising inventions that have yet to reach their full potential – such as the ever-fascinating Sophia.

References

- Anthony, L. (2022). *AntConc* (Version 4.0.5) [Computer Software]. Tokyo, Japan: Waseda University. Available from <https://www.laurenceanthony.net/software>
- Arun, R., Suresh, V., Madhavan, C. V., & Murthy, M. N. (2010, June). On Finding the Natural Number of Topics with Latent Dirichlet Allocation: Some Observations. In *Pacific-Asia conference on knowledge discovery and data mining* (pp. 391-402). Springer, Berlin, Heidelberg.
- Bail, C. (2019, May 7). *Topic Modeling*. The Summer Institutes in Computational Social Science. Retrieved April 11, 2022, from https://sicss.io/2019/materials/day3-text-analysis/topic-modeling/rmarkdown/Topic_Modeling.html
- Berg, A., Buffie, E. F., & Zanna, L. F. (2016). Robots, growth, and inequality. *Finance & Development*, 53(3), 10-13.
- Blom, J. N., Rønlev, R., Hansen, K. R., & Ljungdalh, A. K. (2021). The potentials and pitfalls of interactional speculations by journalists and experts in the media: The case of Covid-19. *Journalism Studies*, 22(9), 1142-1160.
- Bonnington, C. (2021, May 22). *Facebook's AI chief doesn't like Sophia the Robot*. The Daily Dot. Retrieved April 24, 2022, from <https://www.dailydot.com/debug/facebook-ai-chief-sophia-the-robot/>
- Broadbent, E. (2017). Interactions with robots: The truths we reveal about ourselves. *Annual review of psychology*, 68, 627-652.
- Bruckenberger, U., Weiss, A., Mirnig, N., Strasser, E., Stadler, S., & Tscheligi, M. (2013, October). The good, the bad, the weird: Audience evaluation of a "real" robot in relation to science fiction and mass media. In *International conference on social robotics* (pp. 301-310). Springer, Cham.
- Cave, S., Craig, C., Dihal, K., Dillon, S., Montgomery, J., Singler, B., & Taylor, L. (2018). *Portrayals and perceptions of AI and why they matter*. The Royal Society.
- Chan, K. (2018, January 16). *Lifelike robots made in Hong Kong meant to win over humans*. AP News. Retrieved April 25, 2022, from <https://apnews.com/article/f8298725cc6941db8b15c9252060ea76>
- Chella, A., Cangelosi, A., Metta, G., & Bringsjord, S. (2019). Consciousness in humanoid robots. *Frontiers in Robotics and AI*, 6, 17.
- Chong, D., & Druckman, J. N. (2007). Framing theory. *Annual Review of Political Science*, 10, 103-126
- Dafonte-Gómez, A. (2018). News Media and the Emotional Public Sphere | Audiences as Medium: Motivations and Emotions in News Sharing. *International journal of communication*, 12, 20.
- Döring, N., & Poeschl, S. (2019). Love and sex with robots: a content analysis of media representations. *International Journal of Social Robotics*, 11(4), 665-677.
- Fairclough, N. (2013). *Critical discourse analysis: The critical study of language*. Routledge.
- Fortunati, L., Sorrentino, A., Fiorini, L., & Cavallo, F. (2021). The rise of the roboid. *International Journal of Social Robotics*, 13(6), 1457-1471.
- Ghosh, S. (2018, November 26). *Facebook's AI boss described Sophia the robot as "complete b-----t" and "Wizard-of-Oz AI."* Business Insider. Retrieved December 17,

- 2021, from <https://www.businessinsider.com/facebook-ai-yann-lecun-sophia-robot-bullshit-2018-1?international=true&r=US&IR=T>
- Goode, E., & Ben-Yehuda, N. (2012). Moral panic. In *The Routledge Handbook of Deviant Behavior* (pp. 46-52). Routledge.
- Goode, E., & Ben-Yehuda, N. (1994). Moral panics: Culture, politics, and social construction. *Annual review of sociology*, 20(1), 149-171.
- Griffiths, S. (2016, March 17). *Humanoid called Sophia reveals "her aspirations" in an interview with her maker*. Mail Online. Retrieved April 22, 2022, from <https://www.dailymail.co.uk/sciencetech/article-3496982/Eerily-lifelike-robots-walk-just-20-years-don-t-destroy-Humanoid-called-Sophia-reveals-aspirations-interview-maker.html#:~:text=Eerily%20lifelike%20robots%20will%20walk,an%20interview%20with%20her%20maker>
- Hanson Robotics Limited. (2018, September 5). *Sophia the Robot by Hanson Robotics* [Video]. YouTube. Retrieved October 3, 2022, from <https://www.youtube.com/watch?v=BhU9hOo5Cuc>
- Ho, C. C., & MacDorman, K. F. (2010). Revisiting the uncanny valley theory: Developing and validating an alternative to the Godspeed indices. *Computers in Human Behavior*, 26(6), 1508-1518.
- Howell, P. (2016, March 25). *Do androids dream of getting an Oscar?* The Star. Retrieved April 25, 2022, from <https://www.thestar.com/entertainment/movies/2016/03/25/do-androids-dream-of-getting-an-oscar.html>
- Hu, Y., Boyd-Graber, J., Satinoff, B., & Smith, A. (2014). Interactive topic modeling. *Machine Learning*, 95(3), 423-469.
- Ireton, C., & Posetti, J. (2018). *Journalism, fake news & disinformation: handbook for journalism education and training*. Unesco Publishing.
- Jänicke, S., Franzini, G., Cheema, M. F., & Scheuermann, G. (2015, May). On Close and Distant Reading in Digital Humanities: A Survey and Future Challenges. In *EuroVis (STARs)* (pp. 83-103).
- Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y., & Zhao, L. (2019). Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey. *Multimedia Tools and Applications*, 78(11), 15169-15211.
- Jordan, J. M. (2016). *Robots*. MIT Press.
- Jørgensen, M. W., & Phillips, L. J. (2002). *Discourse analysis as theory and method*. Sage.
- Jost, J. T., Banaji, M. R., & Nosek, B. A. (2004). A decade of system justification theory: Accumulated evidence of conscious and unconscious bolstering of the status quo. *Political psychology*, 25(6), 881-919.
- Kaplan, A. D., Sanders, T., & Hancock, P. A. (2019). The relationship between extroversion and the tendency to anthropomorphize robots: A Bayesian analysis. *Frontiers in Robotics and AI*, 5, 135.
- Maier, D., Waldherr, A., Miltner, P., Wiedemann, G., Niekler, A., Keinert, A., ... & Adam, S. (2018). Applying LDA topic modeling in communication research: Toward a valid and reliable methodology. *Communication Methods and Measures*, 12(2-3), 93-118.
- Majid, A. (2022, September 30). *Top 50 largest news websites in the world: Sputnik, Drudge and Fox see biggest traffic falls in February*. Press Gazette. Retrieved November 24,

- 2022, from <https://pressgazette.co.uk/news/top-50-largest-news-websites-in-the-world-sputnik-drudge-and-fox-see-biggest-traffic-falls-in-february/>
- Martin, S. (2016, October 25). *Rise of AI as incredibly lifelike humanoid interacts and tells jokes with humans*. Express. Retrieved April 24, 2022, from <https://www.express.co.uk/news/science/725087/Rise-of-AI-lifelike-humanoid-interacts-with-humans>
- Mayorga, M. W., Hester, E. B., Helsel, E., Ivanov, B., Sellnow, T. L., Slovic, P., ... & Frakes, D. (2020). Enhancing Public Resistance to “Fake News” A Review of the Problem and Strategic Solutions. *The handbook of applied communication research*, 197-212.
- Meinecke, L., & Voss, L. (2018). I robot, you unemployed: robotics in science fiction and media discourse. Schafft Wissen. *Gemeinsames und geteiltes Wissen in Wissenschaft und Technik*, 2, 203-221.
- Mubin, O., Wadibhasme, K., Jordan, P., & Obaid, M. (2019). Reflecting on the presence of science fiction robots in computing literature. *ACM Transactions on Human-Robot Interaction (THRI)*, 8(1), 1-25.
- Müller, B. C., Gao, X., Nijssen, S. R., & Damen, T. G. (2020). I, Robot: How human appearance and mind attribution relate to the perceived danger of robots. *International Journal of Social Robotics*, 1-11.
- Northern Advocate. (2020, September 22). *World’s most human robot hints at future on Northland visit*. NZ Herald. Retrieved April 22, 2022, from <https://www.nzherald.co.nz/northern-advocate/news/worlds-most-human-robot-hints-at-future-on-northland-visit/IT6WIDQ33R7U6HCGXSVAZW316U/>
- Paladino, M. P., Vaes, J., & Jetten, J. (2021). Motivational and Cognitive Underpinnings of Fear of Social Robots That Become “Too Human for Us”. In *The Routledge handbook of dehumanization* (pp. 292-305). Routledge.
- Peters, K., Kashima, Y., & Clark, A. (2009). Talking about others: Emotionality and the dissemination of social information. *European Journal of Social Psychology*, 39(2), 207–222.
- Pettit, H. (2018, January 18). *Facebook AI boss slams humanoid robot Sophia*. Mail Online. Retrieved April 24, 2022, from <https://www.dailymail.co.uk/sciencetech/article-5284951/Facebook-AI-boss-slams-humanoid-robot-Sophia.html>
- Philo, G. (2014). *Seeing and believing: The influence of television*. Routledge.
- Poulter, S., & Pettit, H. (2018, January 9). *World’s first robot citizen “Sophia” gets her own legs*. Mail Online. Retrieved April 25, 2022, from <https://www.dailymail.co.uk/sciencetech/article-5251185/Worlds-robot-citizen-Sophia-gets-legs.html>
- Prasad, R., Saruwatari, H., & Shikano, K. (2004). Robots that can hear, understand and talk. *Advanced Robotics*, 18(5), 533-564.
- Rhee, J. (2018). *The Robotic Imaginary: The Human and the Price of Dehumanized Labor*. U of Minnesota Press.
- Riccio, T. (2021). Sophia Robot: An Emergent Ethnography. *TDR*, 65(3), 42-77.
- Rimé, B. (2009). Emotion elicits the social sharing of emotion: Theory and empirical review. *Emotion Review*, 1(1), 60–85.

- Rocklage, M. D., Rucker, D. D., & Nordgren, L. F. (2018). Persuasion, emotion, and language: The intent to persuade transforms language via emotionality. *Psychological science*, 29(5), 749-760.
- Salgado, S., & Strömbäck, J. (2012). Interpretive journalism: A review of concepts, operationalizations and key findings. *Journalism*, 13(2), 144-161.
- Sardar, A., Joosse, M., Weiss, A., & Evers, V. (2012, March). Don't stand so close to me: users' attitudinal and behavioral responses to personal space invasion by robots. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction* (pp. 229-230).
- Savela, N., Turja, T., Latikka, R., & Oksanen, A. (2021). Media effects on the perceptions of robots. *Human Behavior and Emerging Technologies*, 1-15.
- Scheufele, D. A., & Iyengar, S. (2012). The state of framing research: A call for new directions. *The Oxford handbook of political communication theories*, 1-26.
- Scheutz, M. (2012). The Inherent Dangers of Unidirectional Emotional Bonds Between Humans and Social Robots. In P. Lin, K. Abney, & G. A. Bekey (Eds.), *Robot Ethics: The Ethical and Social Implications of Robotics* (pp. 205-221). MIT Press.
- Schütze, H., Manning, C. D., & Raghavan, P. (2008). *Introduction to information retrieval* (Vol. 39, pp. 234-265). Cambridge: Cambridge University Press.
- Schwartz, O. (2018, October 31). "The discourse is unhinged": how the media gets AI alarmingly wrong. *The Guardian*. Retrieved September 28, 2022, from <https://www.theguardian.com/technology/2018/jul/25/ai-artificial-intelligence-social-media-bots-wrong>
- Sharkey, N. (2018, November 17). *Mama Mia It's Sophia: A Show Robot Or Dangerous Platform To Mislead?* *Forbes*. Retrieved April 24, 2022, from <https://www.forbes.com/sites/noelsharkey/2018/11/17/mama-mia-its-sophia-a-show-robot-or-dangerous-platform-to-mislead/?sh=5f36f49d7ac9>
- Shaw, J. J. (2015). From homo economicus to homo roboticus: an exploration of the transformative impact of the technological imaginary. *International Journal of Law in Context*, 11(3), 245-264.
- Simons, R. (2019, November 21). *Holly Willoughby and Phillip Schofield are creeped out by world's most intelligent robot Sophia*. *Mail Online*. Retrieved April 24, 2022, from <https://www.dailymail.co.uk/tvshowbiz/article-7710419/Holly-Willoughby-Phillip-Schofield-creeped-worlds-intelligent-robot-Sophia.html>
- Søraa, R. A. (2019). Mecha-media: How are androids, cyborgs, and robots presented and received through the media? In *Rapid automation: concepts, methodologies, tools, and applications* (pp. 12-30). IGI Global.
- Steger, M. B., & James, P. (2013). Levels of subjective globalization: Ideologies, imaginaries, ontologies. *Perspectives on Global Development and Technology*, 12(1-2), 17-40.
- Stubbs, M. (1994). Grammar, text, and ideology: computer-assisted methods in the linguistics of representation. *Applied linguistics*, 15(2), 201-223.
- Suchman, L. A. (2007). *Human-Machine Reconfigurations: Plans and Situated Actions* (2nd ed.). Cambridge University Press.
- Swedberg, R. (2020). Exploratory research. *The production of knowledge: Enhancing progress in social science*, 17-41.

- Törnberg, A., & Törnberg, P. (2016). Muslims in social media discourse: Combining topic modeling and critical discourse analysis. *Discourse, Context & Media*, 13, 132-142.
- Urbi, J., & Sigalos, M. (2018, June 5). *The complicated truth about Sophia the robot — an almost human robot or a PR stunt*. CNBC. Retrieved December 17, 2021, from <https://www.cnn.com/2018/06/05/hanson-robotics-sophia-the-robot-pr-stunt-artificial-intelligence.html>
- Vincent, J. (2019, March 27). *Turing Award 2018: Nobel Prize of computing given to 'godfathers of AI.'* The Verge. Retrieved December 17, 2021, from <https://www.theverge.com/2019/3/27/18280665/ai-godfathers-turing-award-2018-yoshua-bengio-geoffrey-hinton-yann-lecun>
- Wallach, H. M. (2006, June). Topic modeling: beyond bag-of-words. In *Proceedings of the 23rd international conference on Machine learning* (pp. 977-984).
- Weston, P. (2017, June 8). *"AI is good for the world" insists Sophia humanoid robot*. Mail Online. Retrieved April 24, 2022, from <https://www.dailymail.co.uk/sciencetech/article-4584770/AI-good-world-insists-Sophia-humanoid-robot.html>
- Wishart-Smith, H. (2021, September 23). *With Feeling: Innovation, Creativity And Empathic Machines*. Forbes. Retrieved April 24, 2022, from <https://www.forbes.com/sites/heatherwishartsmith/2021/08/16/with-feeling-innovation-creativity-and-empathic-machines/?sh=7dbd8f253113>
- Wootson, C. R., Jr. (2017, October 29). *Saudi Arabia, which denies women equal rights, makes a robot a citizen*. Washington Post. Retrieved April 25, 2022, from <https://www.washingtonpost.com/news/innovations/wp/2017/10/29/saudi-arabia-which-denies-women-equal-rights-makes-a-robot-a-citizen/>

Appendix A: Code Used in R (ver. 4.1.3)

Note. This code is an adaptation of the topic modeling guide that is available through <https://ldavis.cpsievert.me/reviews/reviews.html>. The code was adjusted when necessary; several lines were added or altered, such as the value of k and the pre-processing steps.

```

1. # opening the corpus
2. library(readxl)
3. corpussophia <- read_excel("filename.xlsx")
4. dfdata = data.frame(corpussophia)
5. View(dfdata)
6.
7. # importing a list of stopwords
8. library(tm)
9. stop_words <- stopwords("SMART")
10.
11. # pre-processing
12. dfdata <- tolower(as.matrix(dfdata)) # force to lowercase
13. library(textstem)
14. dfdata <- lemmatize_strings(as.matrix(dfdata)) # lemmatization
15. dfdata <- gsub("'|'|\"|'|-|-", "", as.matrix(dfdata)) # remove punctuation
16. dfdata <- gsub("[[:punct:]]", " ", as.matrix(dfdata)) # remove punctuation
17. dfdata <- gsub("[[:digit:]]+", "", as.matrix(dfdata)) # remove numbers
18. dfdata <- gsub("[[:cntrl:]]", " ", as.matrix(dfdata)) # remove control chars
19. dfdata <- gsub("^[:space:]+", "", as.matrix(dfdata)) # remove whitespaces
20. dfdata <- gsub("[[:space:]]+$", "", as.matrix(dfdata)) # remove whitespaces
21.
22. # tokenizing
23. doc.list <- strsplit(dfdata, "[[:space:]]+")
24.
25. # computing the terms table
26. term.table <- table(unlist(doc.list))
27. term.table <- sort(term.table, decreasing = TRUE)
28.
29. # removing stopwords and terms that appear infrequently
30. del <- names(term.table) %in% stop_words | term.table < 5
31. term.table <- term.table[!del]
32. vocab <- names(term.table)
33.
34. # preparing the data for the lda package
35. get.terms <- function(x) {
36.   index <- match(x, vocab)
37.   index <- index[!is.na(index)]
38.   rbind(as.integer(index - 1), as.integer(rep(1, length(index))))
39. }
40. documents <- lapply(doc.list, get.terms)
41.
42. # computing statistics related to the data set
43. D <- length(documents) # number of documents
44. W <- length(vocab) # number of terms in the vocab
45. doc.length <- sapply(documents, function(x) sum(x[2, ])) # number of tokens per
  document
46. N <- sum(doc.length) # total number of tokens in the data
47. term.frequency <- as.integer(term.table) # frequencies of terms in the corpus
48.
49. # MCMC and model tuning parameters
50. K <- 22
51. G <- 5000
52. alpha <- 0.02
53. eta <- 0.02
54.
55. # fitting the model
56. library(lda)

```

```
57. fit <- lda.collapsed.gibbs.sampler(documents = documents, K = K, vocab = vocab,
58.                                   num.iterations = G, alpha = alpha,
59.                                   eta = eta, initial = NULL, burnin = 0,
60.                                   compute.log.likelihood = TRUE)
61.
62. theta <- t(apply(fit$document_sums + alpha, 2, function(x) x/sum(x)))
63. phi <- t(apply(t(fit$topics) + eta, 2, function(x) x/sum(x)))
64.
65. Sophia <- list(phi = phi,
66.               theta = theta,
67.               doc.length = doc.length,
68.               vocab = vocab,
69.               term.frequency = term.frequency)
70.
71. library(LDAvis)
72.
73. # creating the JSON object to feed the visualization
74. json <- createJSON(phi = Sophia$phi,
75.                   theta = Sophia$theta,
76.                   doc.length = Sophia$doc.length,
77.                   vocab = Sophia$vocab,
78.                   term.frequency = Sophia$term.frequency,
79.                   R = 50)
80.
81. serVis(json, out.dir = 'vis', open.browser = TRUE)
```

Appendix B: News Sources Included in Corpus

- ABC News
- AP News
- BBC News
- Bloomberg
- CBS News
- CNN
- Daily Express
- Daily Mail
- Forbes
- HuffPost
- Los Angeles Times
- National Post
- NBC News
- New York Post
- Newsweek
- NZ Herald
- Reuters
- The Globe and Mail
- The Guardian
- The New York Times
- The Wall Street Journal
- The Washington Informer
- The Washington Post
- Toronto Star
- US News
- USA Today

Appendix C: Overview of Generated Topics

1: robot human question create humanoid understand answer life people conversation give robotic research year develop part development world real idea play turn find mind computer face hold form put learn response important base feel end report field accord complex time sort day work platform follow role true numb fear lot

2: robotic ben vega scientist goertzel hanson uncanny professor kong hong las show park valley university speech robot blink theme trade science intelligence electronic consumer ces movement Sophia leg eventually service walk chief mobile refine therapy machine software copy vacuum wow friendly challenge customer team live build black company coffee wheel

3: saudi woman arabia citizenship country grant family male guardian kingdom riyadh female citizen public foreign law abaya initiative guardianship week daughter nation cover sophia wear investment user headscarf obtain saudia prince city group lift hashtag hijab muslim strict child time blood call arabian abroad arabic bin crown foreigner gulf khaleej

4: great human experience machine good team thing social communication moment present bias increase agent interface attention positive personal full opportunity diversity past creative brain test define mirror verbal cognitive communicate mit responsibility theyre workplace extend dont biological cooperate desire inclusion successful inspire integrate kind impact multiple write increasingly imagine

5: hanson camera realistic include sophia eye facial expression make action algorithm expressive model intend sophias creation feature simulate skin face im wife smile audrey hansons contact capable intrigue human david classic recently color embody transformer creator beauty hepburn sense nose natural hope happen speech equip connection exist creativity cheekbone foster

6: decision business organization result network drive learn system focus improve maana ai reason tool knowledge good similar datum performance today hand expertise recommendation support train neural individual current death type online group neuron path quality thompson interaction technology language commercial decide influence constraint machine prove domain require goal large advance

7: pandemic covid mass grace healthcare executive coronavirus hong kong robot elderly robotic price isolate pepper reuters market roll thousand decrease hospital manufacturer softbank unit manufacture launch china produce social health socially plan airline asian autism hoorn realise retail hanson ten company diagnose graces sale accelerate aim prototype lonely korea interaction

8: job economy macdonald technology productivity zealand work automate fund werner industrial smart automation generation benefit analysis sector energy man future task book military replace effective farmer presence workforce economic general involve age vehicle

technological efficient labour tax worker government car primary growth profit taxi pick
conclude daily employment propaganda review

9: mr picture impressive consciousness presenter worlds insist patient con morning pro
negative morgan britain revolutionise mischievously viewer super hanson david fully sophia
television pair bizarre dance outweigh reid someday intelligence teacher afp arent elder
endlessly susanna acknowledge expect resource ultimately ethical sentient discuss rapidly
wise clear concern care consequence terminator

10: blade runner disney texas einstein lim interactive partly revenue automaton entertainment
gimmick motion theatrics print flesh quickly south albert educational jeanne predecessor
sculpt sculptor startup dick power domestic masterpiece profitable shoebox synthesis
problem process genius combine arm win artificial character fine entertain billion dozen sci
progress autistic generally serve background

11: ai weapon nuclear humanity warn school promise pose policy state potential scale treaty
fast power security industry international lie lead driverless killer major artificial agree war
sign guarantee policymaker privacy union intelligence danger technical threat acquire
champion proliferation pull hand debate risk public survey technology issue elon publish
control force

12: artwork nft art digital artist auction bonaceto paint portrait sell nfts fungible physical
collaboration gateway nifty andrea instantiation italian painting sale collector token piece
sophia ownership buyer pop work collaborate image drawing gallery original signature verify
unique completely million element thursday ledger item silver generate title assistance
surface musician blockchain

13: agi singularitynet ai company datum blockchain developer decentralize big token product
user invest marketplace popular store contract pseudo ais solution project segment access
goertzel system dr source market common oasis philosophical service deliver song provide
intelligence customer culture ocean accomplish attract buy schedule ensure list contribute
protocol collect narrow reputation

14: sophia puppet asimo lecun profile westinghouse side series move honda grant facebook
electro inform yann line publicity body fashion mechanical early drone clunky deceive lecuns
televox wensley intelligent appearance step citizenship post strike celebrity october content
follower back represent leg december sophisticate month history make pretend hype write
remarkable add

15: sorkin elon nice sophia musk happy proud ross distinction hollywood future honor youve
youre andrew ill bbc historical input citizenship watch live movie good dig worry respond
sophias read interview quick historic regard tesla dont robot wednesday honour destroy bad
emotion human prevent tweet unique mock output build smart crowd

16: fallon jimmy fake news tonight compassion isnt ai behead rock scissor wouldnt fail release controversy inventor apocalypse feeling ahmed irony newsweek realize audience account afraid cnbc beat article fundamental musk press recognition house receive paper headline civilization claim notably population game human commit dominate president medium prediction short fashion solve

17: hanson camera realistic include sophia eye facial expression make action algorithm expressive model intend sophias creation feature simulate skin face im wife smile audrey hansons contact capable intrigue human david classic recently color embody transformer creator beauty hepburn sense nose natural hope happen speech equip connection exist creativity cheekbone

18: android marry entity civil treat pleo roboticist anxiety corporation treatment turing dr spot surpass demand animal lifelike age complexity hanson live society maturity tip vote intelligent legal pass peoples feel ethical enter class land awaken feeling sci machine include lawmaker study caution election push attempt law generally deserve free fi

19: date smith fan actor teigen atla holly island fly star don friend phenomenon music instagram literally phillip sea zone bite youtube unsettle sophia scary disturb cool air concern lean scare cut wednesday list philip box enjoy career feature twitter video woo photo post encounter didnt robot thursday fight apparently movie

20: rise charlie hawk moore stephen scientist teach wait spell theory super intelligence figure race minute perception andrew artificial everyday threat complicate intelligent computer thing goal interview relate benevolent line wisdom safety sophia instance care damage eventually david multiple compassionate possibility existential check predict company facial critic environmental explain smart hanson

21: conference event soul sofia torso johns reflect tech designer british speaker global vision han attend discuss year future minister dubai laugh audience ceo founder aim goal debate destroy magazine front shoot malleable technological cnbc compute business years conscious prime visitor theyll annual extraordinary dream investment reality program image spread fiction

22: asan azerbaijan mep visa aliyev european baku status agency legal electronic arrival commission liable kill person inappropriate airport government letter issue damage eu service standard parliament president statement ethical centre clean establish perspective visit capacity tour international deliver lead stop safety lawmaker special specialist innovative meet author dilemma car perception

Appendix D: Interpretation of Relevant Topics

1: In my interpretation, topic one is related to the idea that robots can become cognizant beings; it includes words such as “robot”, “humanoid”, “life”, “understand”, “create”, “mind”, “learn”, “real”, and “idea”.

3: Topic three is Sophia’s legally recognized status as a citizen in Saudi Arabia, including words such as “saudi”, “arabia”, “woman”, “grant”, “citizenship”, “foreign”, and “law”.

4: Topic four is related to robots being seen as positive social agents. It includes words such as “great”, “experience”, “machine”, “team”, “communication”, “positive”, “opportunity”, “diversity”, “kind”, and “inspire”.

5: Topic five is related to Sophia’s physical appearance and capabilities, including words such as “realistic”, “eye”, “camera”, “speech”, “facial”, “expression”, “skin”, and “smile”.

7: Topic seven is about Sophia’s potential role in healthcare, featuring words such as “pandemic”, “healthcare”, “elderly”, “lonely”, “isolate”, “hospital”, and “diagnose”.

8: Topic eight is about the role of automation in the economy. It includes words such as “job”, “economy”, “productivity”, “automate”, “industrial”, “smart”, “benefit”, “replace”, “worker”, and “effective”.

9: Topic nine is related to Sophia’s appearance on Good Morning Britain, featuring words such as “presenter”, “morning”, “morgan”, “Britain”, “viewer”, “television”, “reid”, and “susanna” – “morgan”, “reid”, and “susanna” refer to presenters Piers Morgan and Susanna Reid.

10: Topic ten is entertainment-related, and encompasses one of my expectations: that science fiction plays a role in news coverage of robotics. This topic includes words like “blade”, “runner”, “disney”, “entertainment”, “gimmick”, “theatrics”, “automaton”, “character”, and “power”.

11: Topic eleven has to do with the role of automation in political conflict, including words such as “ai”, “weapon”, “nuclear”, “war”, “policy”, “power”, “security”, “killer”, “driverless”, “danger”, and “threat”.

14: Topic fourteen is computer scientist Yann LeCun’s criticism of Sophia. This topic includes words such as “yann”, “lecun”, “sophia”, “puppet”, “pretend”, “hype”, “inform”, “publicity”, “clunky”, and “deceive”.

16: Topic sixteen is about Sophia’s presence in the public eye, such as on late-night talk shows. This topic includes words such as “jimmy”, “fallon”, “tonight”, “controversy”, “newsweek”, “fashion”, “press”, and “recognition”. Interestingly, the topic also includes the

words “game”, “rock”, “paper”, and “scissor” – most likely due to the game of rock, paper, scissor that Sophia played with Jimmy Fallon on The Tonight Show.

18: Topic eighteen appears to be related to the perception of androids as civil entities and the unease this causes, including words such as “android”, “civil”, “entity”, “marry”, “anxiety”, “lifelike”, “society”, “legal”, “ethical”, “awaken”, “caution”, “sci”, and “fi”.

19: Finally, topic nineteen relates to Sophia’s presence in online culture. The topic includes words such as “fan”, “phenomenon”, “instagram”, “youtube”, “twitter”, “career”, “movie”, and “feature”. What is particularly interesting about this topic is that it also includes words such as “scary”, “disturb”, “concern”, and “unsettle”.