OPEN FORUM

AI in the headlines: the portrayal of the ethical issues of artificial intelligence in the media

Leila Ouchchy[1] · Allen Coin[1] · Veljko Dubljević[1]

Abstract

As artificial intelligence (AI) technologies become increasingly prominent in our daily lives, media coverage of the ethical considerations of these technologies has followed suit. Since previous research has shown that media coverage can drive public discourse about novel technologies, studying how the ethical issues of AI are portrayed in the media may lead to greater insight into the potential ramifications of this public discourse, particularly with regard to development and regulation of AI. This paper expands upon previous research by systematically analyzing and categorizing the media portrayal of the ethical issues of AI to better understand how media coverage of these issues may shape public debate about AI. Our results suggest that the media has a fairly realistic and practical focus in its coverage of the ethics of AI, but that the coverage is still shallow. A multifaceted approach to handling the social, ethical and policy issues of AI technology is needed, including increasing the accessibility of correct information to the public in the form of fact sheets and ethical value statements on trusted webpages (e.g., government agencies), collaboration and inclusion of ethics and AI experts in both research and public debate, and consistent government policies or regulatory frameworks for AI technology.

1 Introduction

The practical applications and prevalence of artificial intelligence (AI) in our daily lives has grown significantly in recent years (Standing Committee of the One Hundred Year Study of Artificial Intelligence 2016). With this increase in prevalence and applicability of AI has come a wide range of ethical debates, including how AI can be programmed to make moral decisions, how these decision-making processes can be made sufficiently transparent to humans, and who should be held accountable for these decisions.

A crucial objective in the development of ethical AI is cultivating public trust and acceptance of AI technologies.

In the academic literature, concerns have been voiced about the business drivers for monetizing AI (see Baldini et al. 2018; Bauer and Dubljević 2019), especially since the most successful strategies for AI development (bottom-up approaches and, to a certain extent, hybrid approaches) and monetization create a lack of transparency. Additionally, the media has a large impact on the way issues are framed to the public in every field (Racine et al. 2005; Royal Society 2018; Chuan et al. 2019), so public opinion and acceptance of AI will likely be impacted by the media they consume about it. Because the members of the general public, as both consumers in the market economy and constituents of a liberal democracy, are key stakeholders for technology adoption—and, to a certain extent, for public policy and regulatory oversight—public opinion could affect what kind of AI is developed in the future and how AI is regulated by the government. For these reasons, it is important to analyze how issues of AI and ethics are portrayed in the media. In certain instances where a topic is highly stigmatized, there is a marked disconnect between media representation and public opinion (see Ding 2009), but there are no initial reasons to suspect that would be the case with AI ethics. Going

✉ Veljko Dubljević
veljko_dubljevic@ncsu.edu

Leila Ouchchy
louchch@ncsu.edu

Allen Coin
allen_coin@ncsu.edu

1 Department of Philosophy and Religious Studies, North Carolina State University, 101 Lampe Drive, Raleigh, NC, USA

forward, it is important to gather more data and ascertain if a mismatch in public expectations is actually occurring.

Previous work on the media representation of other novel technologies has emphasized the disruptive potential of overly enthusiastic media coverage to distort the publicly available information (see Dubljević et al. 2014). Additionally, lack of engagement in the public discussion by experts and informed decision makers may lead to a polarization effect in the public, where "hype and hope" and "gloom and doom" perspectives distort the debate (Dubljević 2012, p. 69; see also Oren and Petro 2004). These perspectives can also stall implementation of effective policies, which could be used to increase the public benefit of novel technologies and reduce their harmful effects for individuals and society.

Until recently, narratives about AI have belonged to the realm of fiction. These portrayals have been characterized as extreme, with representations that are "either exaggeratedly optimistic about what the technology might achieve, or melodramatically pessimistic" (Royal Society 2018, p. 9). In recent years, however, news stories have emerged detailing the tangible developments of AI. This has prompted discussion on the effect that narratives, both real and fictional, have on the public perception of AI. In their paper entitled "Framing Artificial Intelligence in American Newspapers," Chuan et al. (2019) conducted a literature review on the portrayal of AI in American news sources. Although they report that the "topic of ethics has been increasingly discussed in recent years," they also found that most of the articles in their sample "did not discuss a particular ethical issue in-depth, but raised general questions about potential ethical concerns" (p. 343).

In 2016, Stahl and colleagues published a paper which provided the first comprehensive systematic review of the academic literature on the ethics of computing (Stahl et al. 2016). This paper examined academic publications published between 2002 and 2012 having to do with the ethics of computing. The authors of this paper found that a wide range of topics related to ethics of computing were covered, and that there was focused coverage of many different technologies. However, they also concluded that discussion of the methodologies, recommendations, and contributions of papers was lacking (Stahl et al. 2016).

To draw on and add to the discussion, we have conducted a review of English language media sources covering the ethics of AI. Unlike the review conducted by Chuan et al., we chose to include only articles covering the ethics of AI because of the salience these articles have for public opinion. Additionally, we review all of the relevant articles as opposed to attempting to create a representative sample, since our focus is on media as a whole as opposed to articles from specific publishers (cf. Chuan et al. 2019). Moreover,

we give special attention to the recommendations given by these articles in relation to ethical issues raised.

The paper by Stahl and colleagues focused on ethical issues, but provided an overview of the ethics of computing technologies in general, whereas ours will focus specifically on ethical issues in artificial intelligence. Because of the impact the media representation of new technologies has on the public discourse, and the fact that the debate around AI ethics has moved from an academic context to a public one (arguably facilitated by media coverage), we analyze publicly accessible news media sources rather than academic sources. Finally, since the review by Stahl and colleagues only examined literature published through 2012, and found limited discussion of ethics of AI issues, our review will look at literature published since then.

2 Methods

In our analysis of the literature, we used qualitative methods of content analysis to create a comprehensive picture of how the ethics of AI are described to the public. To assemble our sample of articles, we used the NexisUni database, because it includes relevant media sources such as newspapers, magazines and Web blogs. We used the following search terms: "Artificial Intelligence" or "Computational Intelligence" or "Computer Reasoning" or "Computer Vision Systems" or "Computer Knowledge Acquisition" or "Computer Knowledge Representation" or "Machine Intelligence" or "Machine learning" or "Artificial Neural Networks" and (Morals or Moral or Morality or Ethic* or Metaethics).

These terms were designed to collect articles that have the words "Artificial Intelligence" and "Ethics," or similar terms, within two words of each other. The search was conducted on October 11, 2018. Within the search, we also filtered for content type (news), language (English), and date (after 1/1/2013).

This search yielded 903 results. Many of these were duplicates, and with duplicates excluded, the sample contained 479 articles. Out of these 479, we excluded articles that met any of the following exclusion criteria: radio/television transcripts, tangential to our topic, articles covering fictional content, and Web pages without full text.

After excluding 225 articles due to the above criteria, we were left with a final sample of 254 articles. This sample included articles with a range of publication types, including newspaper articles, magazine articles, and blogs. To establish intercoder reliability, two authors (L.O. and V.D.) coded a pilot sample consisting of every tenth article in the

sample when ordered alphabetically. We compared coding and decided on the phrasing and delineations of the different categories we would use. We systematically coded for the following five areas of interest:

- *Issue* ethical issues discussed in the articles relating to the ethics of AI (e.g., privacy).
- *Principles based on ethical frameworks* established ethical theories or principles based on an ethical framework explicitly mentioned in the article (e.g., utilitarianism).

Fig. 1 Levels of coding

- *Recommendation* recommendations given or presented in the article related to the ethics of AI.
- *Tone* the tone of the article regarding AI (enthusiastic, balanced/neutral, or critical).
- *Type of technology* specific AI technologies discussed (e.g., autonomous vehicles).

While the paper by Stahl and colleagues also coded for issue, ethical theory, recommendation, and type; we included tone as a coding area because of its relevance to media samples. Academic sources typically aim to adopt a neutral tone with regard to an issue, but media sources often explicitly take a side and frame issues in certain ways (Entman 1993), which then affects how the general public perceives the technology discussed.

Our first products of coding were raw codes, which we will refer to in this paper simply as "codes." We then organized these raw codes into categories based on their content, sometimes involving multiple levels of categorization (visualized in a flowchart in Fig. 1). Henceforth, we will refer to the largest categories as nodes (Level 1 in Fig. 1) and categories within those nodes as subnodes (Level 2 in Fig. 1). Nodes and subnodes are composed of codes (Level 3 in Fig. 1). After completing our coding and categorization, we gathered our results, which we summarize in the next section.
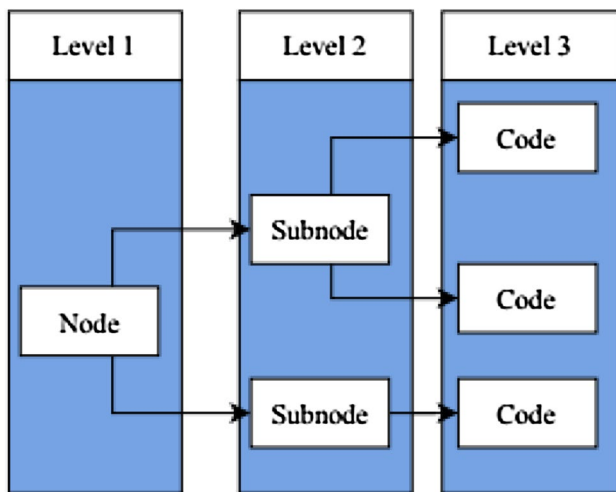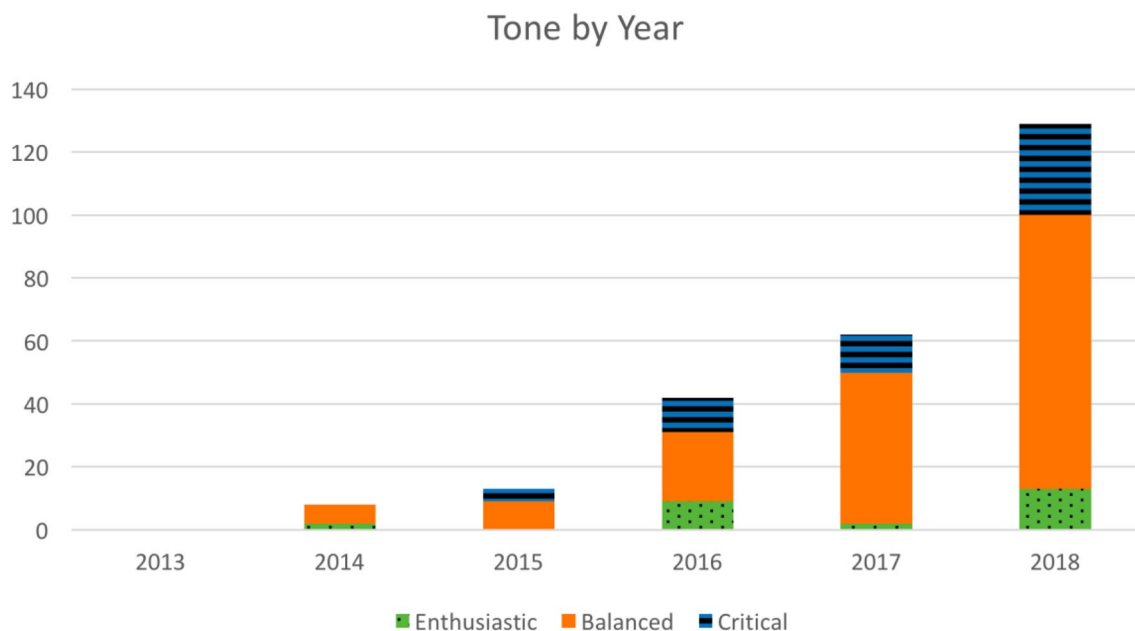

Fig. 2 Tones of articles in different years

3 Results

Although we searched for articles published between the years 2013 and 2018, our final sample included no articles from 2013. The distribution of publication years showed a strong trend toward the later years (see Fig. 2), with the number of articles published in 2013–2017 ($n = 125$) being less than the number of articles published in 2018 alone ($n = 129$). The excluded articles showed the same trend, but had a slightly more balanced distribution.

Our coding revealed that not all of the sources contained all of the areas of interest for which we coded. While all but two of the articles ($n = 252/254$, or 99%) had an identifiable issue, 46 articles (18%) did not have a recommendation. Furthermore, the minority ($n = 28/254$, or 11%) of the articles explicitly mentioned or drew on principles based on ethical frameworks. All of the articles were coded for at least one type of AI technology, but some were coded as "General AI," meaning that they either didn't mention a specific technology, or were covering AI in general. Finally, all articles were coded as having a tone—either enthusiastic, balanced/neutral, or critical.

After coding, we categorized the individual codes according to their subject matter, and created flowcharts to represent the categories. In this section, we will cover the different areas of interest that we coded for, and the results that we found, in order of the coding areas with the greatest to least proportion of articles with codes (see also Table 1 in the Appendix).

3.1 Tone

All articles were coded as having a tone that was either "enthusiastic," "critical," or "balanced/neutral." Enthusiastic articles discussed the benefits of AI from an ethical perspective without discussing the drawbacks, while critical articles discussed the drawbacks without discussing the benefits. Balanced/neutral articles either addressed their issues without a particular tone, or presented both positive and negative aspects of the use of AI.

The majority of the articles were coded as having a balanced/neutral tone, with 173 coded as balanced/neutral, 55 coded as critical, and 26 coded as enthusiastic. Distributions of different tones over the years can be seen in Fig. 2. There was a slight shift from enthusiastic articles to critical articles from 2014 to 2015, but in 2016 the levels were more even. In 2017 and 2018, however, there were significantly more balanced/neutral and critical articles as compared to enthusiastic articles.

3.2 Types of technology

We found that the articles in our sample covered a wide range of types of AI, while many also covered AI in general. Accordingly, we coded the articles based on types of AI that they covered or mentioned (e.g., "Autonomous Vehicles"), as well as "General AI" if their discussion was centered on the topic of AI as a whole. Articles that were coded as only "General AI" for their type discussed AI as a general phenomenon without mentioning any examples of specific types of AI. Articles that were coded as "General AI" as well as other types covered AI generally and also mentioned specific types as examples.

We found 36 different kinds of technologies mentioned, or 37 types total with our "General AI" code. After "General AI," the most common types of technologies mentioned were autonomous vehicles (AVs), autonomous weapons, and military applications. 181 (71%) articles were coded as covering "General AI." Of these, 85 were coded as only "General AI" for their type, whereas 96 were coded as "General AI" among other types. This suggests that the media discussion around AI is less focused on specific types of AI, but will often use specific types of AI as examples in its broader discussion.

3.3 Issues

We coded the articles in our sample according to the issues they discussed with relation to AI and ethics. We found the second highest diversity of individual codes in this section, with 75 different issues coded. Many articles discussed more than one issue, but, as evident by the number of individual codes, there was significant overlap between the issues discussed in different articles.

We organized these codes into the following nine nodes (ordered here by highest to lowest frequency): "undesired results," "accountability," "lack of ethics in AI," "military and law enforcement," "public involvement," "regulation," "human-like AI," "best practices," and "other." In this section, we will discuss the data we collected for the three nodes with the highest frequencies: "undesired results," "accountability," and "lack of ethics in AI."

3.3.1 Undesired results

Articles about the ethics of AI may warn about the dangers of AI going rogue, or other unintended consequences. We used the "undesired results" node to group issues regarding negative outcomes that might arise from the implementation

of AI. This was by far the largest node, both by number of individual codes in it ($n = 33$) and number of occurrences of those codes ($n = 308$). This node made up over half ($n = 308/563$, or 55%) of the total occurrences of issues coded. The subnodes under this node were: "protecting humans from AI," "equity," "economy," "control of AI," "politics," and "human reliance on AI." As with most of the categories, there were the highest number of occurrences in 2018, with over double the number in 2018 compared to 2017.

Our "undesired results" node included three of the top five most common issues coded: "prejudice," "privacy/data protection," and "job loss to AI/economic impact of AI." The most common code, with 57 occurrences, was "prejudice." This often showed up in articles expressing concern over algorithms being biased against certain minority groups. For example, one article that included a discussion about prejudiced AI framed the issue as such:

> Human biases have a way of creeping into code for mass-produced products, giving us automatic soap dispensers that ignore dark skin, digital cameras that confuse East Asian eyes with blinking, surname input fields that reject apostrophes and hyphens, and no shortage of other small indignities… (O'Carroll and Driscoll, The Christian Science Publishing Society, 2018)

This article, as well as 56 other articles in our sample, highlighted the dangers that can come from the reflection, and often magnification, of human biases in AI.

The second most common code in our sample, "privacy/data protection," was also included in this node under the subnode "protecting humans from AI." This code had 55 occurrences. Privacy was the most widely discussed issue in the review published by Stahl and colleagues as well, suggesting that privacy is also a common issue of concern with regards to technology in general, as well as in academic sources.

This node also contained the code: "job loss to AI/economic impact of AI," with 39 occurrences. One might expect articles covering this issue to have an overwhelmingly critical tone, because of its ability to incite fear in readers about maintaining their livelihood. However, this was not the case in our sample: 86% ($n = 33/39$) of articles covering this issue had a balanced/neutral tone.

3.3.2 Accountability

Our "accountability" node included three codes: "transparency," "trust," and "responsibility/liability". Despite only containing three issues, this category comprises 16% ($n = 90/563$) of total occurrences. These occurrences are heavily weighted toward the later years, with more than double the occurrences in 2018 compared to 2017. This may be connected with a spike in mainstream media coverage of unethical practices in the AI sector of the tech industry (see "Discussion").

The most frequent code in this node, and the third most frequent issue code overall, was "transparency," with 45 occurrences. Articles discussing this issue involve concerns over AI technologies making algorithmic decisions that cannot be explained or understood by humans. One article described the issue as such:

> The growth of Artificial Intelligence creates a risk that machine-to-machine decisions could be made without input and transparency to humans. To avoid this, and to ensure AI is developed ethically, individuals need to be able to influence and determine the values, rules, and inputs that guide the development of personalized algorithms directly related to their identity. (Anonymous, Business Wire, 2017

This issue has a strong connection to prejudice, because one of the contributing factors to biased decisions being made by algorithms is a lack of transparency, and 31% ($n = 14/45$) of articles coded for "transparency" were also coded for "prejudice."

3.3.3 Lack of ethics in AI

Some of the articles in our sample discussed specific ethical issues in the creation of AI, as well as the use of ethics in AI. We grouped such articles under the "lack of ethics in AI" node, which had four subnodes: "developing AI," "not enough focus on ethics," "disagreement among humans about ethics," and "violating established ethical guidelines." 13% ($n = 71/563$) of the occurrences of issues were under this node, while it contained 11% ($n = 8/75$) of the individual codes.

3.4 Recommendations

We now turn to another area of coding: recommendations, where we coded articles in our sample based on their recommendations having to do with ethics and AI. This section had the highest diversity of individual codes, with 106 different recommendations coded. We divided the individual recommendations into eleven different nodes: "encouraging public involvement," "avoiding undesired results," "how to regulate AI," "ethics in AI," "research," "implementing best practices," "human in the loop/oversight," "recommendations for military and law enforcement," "choice and responsibility," "anthropomorphizing AI," and "healthcare and AI." In this section, we will discuss the three most common nodes of recommendations found in the literature, and the data we found for them.

3.4.1 Encouraging public involvement

While the "public involvement" node in our Issues section was only the fifth most common node, it was much more common for articles in our sample to give recommendations related to encouraging public involvement. "Encouraging public involvement" was the most frequent recommendation node, with 78 occurrences. This node consists of recommendations having to do with increasing participation and competency in the field of AI from different sectors of the public. This node contains articles from all years except for 2013, but still showed a sharp increase over time, with the number of occurrences in 2018 higher than all of the other years combined. This increase may be due to media coverage of Cambridge Analytica, which drew public attention to the subject of unethical practices in the use of AI technology.

The second most common recommendation code, with 27 occurrences, was "public debate," which was included under this node. This recommendation was present in articles covering a wide variety of issues, including "job loss to AI," "AI decision-making," and "privacy." One article gave the following recommendation:

> If [a driverless car] detects a huge lorry hurtling directly towards it and the only escape is to swerve into a group of pedestrians on the pavement, will it kill or injure them to reduce the likely death toll inside the car?… Although the answers are far from straightforward, they will need to be addressed sooner rather than later, preferably in a public discussion." (Anonymous, Financial Times, 2014)

The prevalence of this recommendation, as well as the "encouraging public involvement" node in general, highlights the importance placed in our sample on large-scale participation in a wide variety of issues surrounding AI.

3.4.2 Avoiding undesired results

70 articles in our sample gave recommendations aimed specifically at avoiding possible negative outcomes of AI, many of which appeared in the "undesired results" section of the issues. We grouped these articles in the node: "avoiding undesired results," which was the second largest node in our recommendations section. It had five subnodes: "anticipatory work," "societal issues," "abuse," "economy/workforce," and "protecting humans from AI." This node had the highest number of codes out of any other recommendations node, with 26 individual codes.

Although the "avoiding undesired results" node contained codes which correspond to the most common issues, such recommendations were not found nearly as frequently as their counterparts in the Issues section. For example, the "avoiding undesired results" node contained two codes related to bias: "incorporate measures to control bias in the models" and "study the cost of fairness with regard to biased AI." However, these codes had only three occurrences together, while the related issue code "prejudice" had 57. This suggests that articles writing about some of the most common issues gave recommendations that were not concretely related to the issue they dealt with.

3.4.3 How to regulate AI

13% ($n=48/373$) of the recommendations coded were grouped under the node "how to regulate AI." We divided this node into four subnodes: "pro-regulation," "against regulation," "encouraged use of AI," and "intellectual property." The "pro-regulation" subnode contained codes encouraging different kinds of regulatory systems for AI, slower deployment of AI, and non-governmental forms of societal intervention. This subnode had 31 occurrences in total, making it the most frequent subnode of this node. The "against regulation" subnode contained three different codes, with five total occurrences. The "encouraged use of AI" subnode, which is related to the "against regulation" subnode because articles that advocate for limited regulation of AI often do so to encourage deployment, had four different codes, with 11 total occurrences. From this node we can see that, although the benefits to faster deployment of AI have been touted (Standing Committee of the One Hundred Year Study of Artificial Intelligence 2016), significantly more articles advocated for governmental involvement and/or regulation of AI than discouraged it.

3.5 Principles based on ethical frameworks

We found 11 sets of principles based on ethical frameworks, which also included some traditional ethical theories, mentioned in the articles. However, the majority of articles (89%, $n=226/254$) did not mention principles based on ethical frameworks. The three most common principles/theories mentioned were Utilitarianism, with eight occurrences; Asimov's "three laws of robotics," with seven occurrences; and the Hippocratic Oath, with five occurrences. Additionally, although 18% ($n=46/254$) of the articles did not have a recommendation, 100% of the articles that mentioned an ethical principle/theory also gave a recommendation.

4 Discussion

The public debate on the ethics of AI is fairly new. For instance, our results show that there were no relevant media articles discussing the issue in 2013, whereas the discussion rapidly increased in the following years. However, it should be noted that it is not expressly the case that the academic debate migrated into the public, but rather that separate highly popularized events, such as the Tesla auto-pilot accident, have also increased public concerns about AI and recognition of the need for ethical oversight. We base this observation on the fact that the vast majority of media articles start with clearly identified ethical issues raised by AI and that most of them offer constructive recommendations, whereas the ethical theories and principles figure only in the minority (i.e., 11%) of the media articles. Additionally, prior research on public opinion has clearly identified widespread concerns about adoption of AI in the USA (Anderson 2017) and the UK (Cave et al. 2019).

4.1 The tone of the AI ethics debate

One of the key issues we sought to understand was whether the media portrayal of emerging ethical issues in AI constitutes an overreaction that could be transferred to the public. As Shariff and colleagues note, "overreactions risk slowing or stalling the adoption of [AI]" (Shariff et al. 2017). In this respect, it should be noted that the public debate on the ethics of AI adequately captures the hopes and fears arising from the rapid introduction of AI technology into society. Initially optimistic and enthusiastic reporting (e.g., in 2014) was followed up by mostly critical or balanced tones in more recent years. As noted in "Results", even articles covering issues that are deeply personal, such as job loss to AI, had overwhelmingly balanced/neutral tones. Thus, unlike media coverage of certain forms of neurotechnology (see Dubljević et al. 2014) where the media coverage was skewed toward the enthusiastic side, or the fictional narratives about AI where the coverage tends to be driven by utopian or dystopian extremes (Royal Society 2018), the public debate around the ethics of AI shows less propensity to be dominated by hype. The reasons for this could be that prior social experience with rapid adoption of technology with potentially disruptive effects on the workforce dictates a more cautious approach, and that there is a healthy dose of appreciation of the charge that AI is "turning many workplaces into human-hostile environments" (Agar 2019, p. 92). Indeed, much in the way that rapid industrialization in prior eras created vast riches for some, while disempowering, dispossessing, and even displacing many others, it is abundantly clear that society needs to be prepared for both positive and negative consequences of AI technology adoption. This partially confirms the findings of the study by Fast and Horvitz (2017), which analyzed 30 years of coverage of AI in the New York Times and reported an overall optimistic view with growing concerns about negative consequences of AI in more recent coverage. Our findings generalize to media coverage beyond a single influential newspaper, qualify the development of the public debate, and focus specifically on ethical issues (as opposed to general issues) of AI.
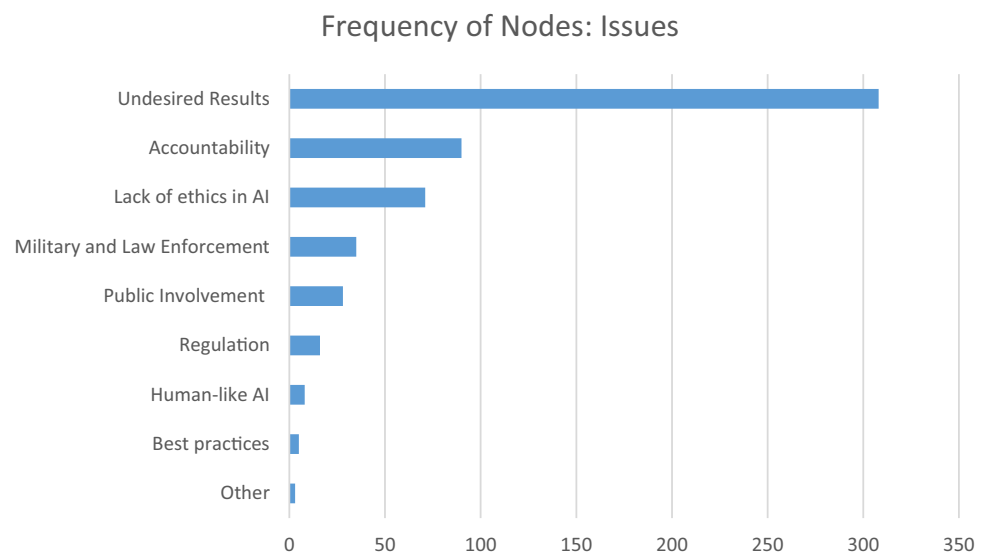
4.2 Issues in the AI ethics debate

The most frequent issues discussed in our sample, such as prejudice and privacy/data protection, were largely practical ones that would have an impact on everyday members of society. When articles did discuss matters of life and death, it was most often in the context of the militarization of AI, which is a reasonable concern. This shows a contrast when compared to television and movie portrayals of AI, which often sensationalize the dangers posed by AI.

The focus on human prejudices being recreated by AI (for example, in hiring algorithms) shows a concern for the effects that emerging technologies will have on marginalized members of society. However, discussion of less overt inequalities, such as who will have access to new technologies, was less prevalent. These findings are in line with the results of Stahl and colleagues (2016) and qualify the findings of Chuan and colleagues (2019). Namely, Stahl and colleagues discuss privacy as the most salient issue in the academic discussion of computing ethics in general, whereas Chuan and colleagues, most likely due to their lack of focus on ethics and the shortcomings of the sampling method, feature the issue of using AI to combat unethical wildlife trade (a topic falling under our criminal justice/ law enforcement code, with 15 occurrences in total), and only mention privacy and the misuse of AI in passing.

4.3 Recommendations in the AI ethics debate

From the most frequent recommendations found in our sample ("developing ethical guidelines" and "public debate,"), it is clear that the public debate on the ethics of AI is in its early stages. However, it appears to be fairly sophisticated, and a crucial recommendation is to increase the breadth and depth of public debate, as well as participation of relevant stakeholders. The increase in media reporting year-over-year signals that the public hopes and concerns about the widespread adoption of AI technology are rapidly felt as a crucial topic for society, which makes it prominent in the

Fig. 3 Ethical issues in AI

Frequency of Nodes: Issues

5 Conclusion

As we noted above, since public opinion is so important to the development and adoption of new technologies, it is crucial to pay attention to how ethical issues of AI are portrayed in the media, and our findings offer a basis for a tentative way forward. The issues most frequently covered, along with the mostly balanced/neutral tones, suggest that the media has a fairly realistic and practical focus in its coverage of the ethics of AI. The media discussion around the ethics of AI is less focused on specific types of AI, but will often use specific examples in its broader discussion. While one might expect articles covering the issue of AI-related unemployment to have an overwhelmingly critical tone, due to its ability to incite fear in readers about maintaining their livelihood, the promise of economic benefits and creation of new jobs is leveraged as a part of a balanced ongoing discussion. Since media articles writing about some of the most common issues gave recommendations that were not concretely related to the issue they dealt with, it is safe to assert that the discussion is sophisticated in tone (e.g., avoiding hype), but not yet in content. This is the case both with the details of AI technology implementation and with the ethical frameworks that are supposed to guide the development. This suggests that the articles about AI and ethics are written by authors with insufficient knowledge of AI technology or ethics.

Therefore, a multifaceted approach to handling the social, ethical and policy issues of AI technology is needed. This could include increasing the accessibility of correct information to the public in the form of fact-sheets and ethical value statements on trusted webpages (e.g., government agencies), to make sure the public is well informed. This could incentivize the media to continue to present an accurate and

agenda-setting stage of the policy process. Our findings align with those of Chuan et al. (2019) in this respect: they found the discussion of AI to be largely stagnant in terms of the number of articles on AI between 2009 and 2013, with a drastic rise starting in 2014 (See Chuan et al. 2019, Fig. 3). The crucial aspect in this nascent, yet sophisticated, public debate is the need to harness potential benefits of AI while at the same time decreasing or mitigating the potential negative effects on society. Thus, the discussion surrounding regulation of AI, whether in the form of self-regulation by industry or through government involvement, is increasing in prominence. While some media articles voiced concerns about regulation stifling the development of AI, the need remains to somehow ensure that potential negative consequences of AI will be avoided, including unemployment and the loss of human lives.

Instances such as the highly publicized fatal Tesla Autopilot accident of 2016, which resulted in the death of 40-year-old Joshua Brown (Singhvi and Russel 2016), have spurred the voices critical of AI implementation and given more credence to pro-regulation approaches. In fact, there is a considerable push toward clear policies that recognize the entitlement of individuals to challenge any decision made by algorithms. At the forefront of these is the General Data Protection Regulation (GDPR), established by the European Union (see EU 2016). In the USA, the tech industry is more or less cognizant of this popular demand for regulation, but so far, they appear to favor the self-regulation model, such as the "Partnership for AI" (Metz 2019), while voices in the academic debate frequently conclude that it is "time for citizens of other democratic societies to request that their rights be protected as well" (Bauer and Dubljević 2019).

balanced portrayal of AI and prevent the potential spread of misinformation on AI in the future. Also, collaboration and inclusion of ethicists and AI experts in both research and public debate could make media coverage of AI more sophisticated in its content, thereby helping to mitigate the lack of certainty, and potentially averting some of the possible undesired social outcomes. Finally, as reflected by the number of recommendations found regarding regulation of AI, a clear government policy or regulatory framework for AI technology in the USA and other non-EU countries is urgently needed.

Compliance with ethical standards

Conflict of interest The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Appendix

See Table 1.

References

Agar N (2019) How to be human in the digital economy. MIT Press, Cambridge, p 231

Anderson M (2017) 6 key findings on how Americans see the rise of automation. Pew Research Center. https://www.pewresearch.org/fact-tank/2017/10/04/6-key-findings-on-how-americans-see-the-rise-of-automation/. Accessed 4 Oct, 2017

Anonymous. Business Wire (2017) IEEE Global Initiative for Ethical Considerations in Artificial Intelligence. Accessed 11 Oct 2018

Anonymous. Financial Times (2014) The Morality of thinking machines: It's time to give artificial intelligence an ethical dimension. Accessed 11 Oct 2018

Baldini G, Botterman M, Neisse R, Tallacchini M (2018) Ethical design in the internet of things. Sci Eng Ethics 24:905–925

Bauer WA, Dubljević V (2019) AI assistants and the paradox of internal automaticity. Neuroethics. https://doi.org/10.1007/s12152-019-09423-6

Cave S, Coughlan K, Dihal K (2019) Scary robots. AAAI workshop: AI, ethics, and society. AAAI Workshops. AAAI Press

Chuan C-H, Tsai W-H, Cho S (2019) Framing artificial intelligence in American newspapers. AAAI workshop: AI, ethics, and society. AAAI workshops. AAAI Press

Ding H (2009) Rhetorics of alternative media in an emerging epidemic: SARS, censorship, and participatory risk communication. Tech Commun Q 18:327–350

Dubljević V (2012) Principles of justice as the basis for public policy on psychopharmacological cognitive enhancement. Law Innov Technol 4(1):67–83

Table 1 The ten most common codes for issues, recommendations, and type

Issues		Recommendations		Types of AI	
Prejudice	57	Developing ethical guidelines	28	General AI	181
Privacy/data protection	55	Public debate	27	Autonomous Vehicles	53
Transparency	45	Transparency/algorithmic transparency	14	Autonomous weapons	30
Job loss to AI/economic impact of AI	39	Collaboration across fields	13	Military applications	23
AI decision-making	35	Empowering ethics committees/ethicists	12	Healthcare applications	22
Responsibility/liability	32	Companies taking responsibility	12	Surveillance AI	16
Weaponization	29	General government involvement/regulation	12	Criminal justice/law enforcement applications	15
AI abuse	27	Human control over AI	12	Risk assessment applications	15
Embedding ethics in AI	24	Considering social impact	11	Chatbots	12
Safety	19	Preparing for changes caused by AI	11	Face recognition	10

Dubljević V, Saigle V, Racine E (2014) The rising tide of tDCS in the media and academic literature. Neuron 82(4):731–736

Entman RM (1993) Framing: toward clarification of a fractured paradigm. J Commun 43(4):51–58

Fast E, Horvitz E (2017) Long-term trends in the public perception of artificial intelligence. AAAI Press, Stanford, pp 963–969

Metz C (2019) Is ethical A.I. even possible? The New York Times. https://www.nytimes.com/2019/03/01/business/ethics-artificial-intelligence.html

O'Carroll E, Driscoll M (2018) 2001: a Space Odyssey turns 50…" The Christian Science Publishing Society. Accessed 11 Oct 2019

Oren TE, Petro P (2004) Global currents: media and technology now. Rutgers University Press, New York, p 273

Racine E, Bar-Ilan O, Illes J (2005) fMRI in the public eye. Nat Rev Neurosci 6(2):159–164

Shariff A, Bonnefon J-F, Rahwan I (2017) Psychological roadblocks to the adoption of self-driving vehicles. Nat Hum Behav. https://doi.org/10.1038/s41562-017-0202-6

Singhvi A, Russel K (2016) Inside the self-driving tesla fatal accident. The New York Times. https://www.nytimes.com/interactive/2016/07/01/business/inside-tesla-accident.html. Accessed 15 Nov 2018

Stahl BC, Timmermans J, Mittelstadt BD (2016) The ethics of computing: a survey of the computing-oriented literature. ACM Comput Surv. https://doi.org/10.1145/2871196

Standing Committee of the One Hundred Year Study of Artificial Intelligence (2016) Artificial intelligence and life in 2030. Technical report. Stanford University, Stanford

The Royal Society (2018) Portrayals and perceptions of AI and why they matter. https://royalsociety.org/-/media/policy/projects/ai-narratives/AI-narratives-workshop-findings.pdf