

“Scary Robots”

Examining Public Responses to AI

Stephen Cave

Leverhulme Centre for the Future
of Intelligence
University of Cambridge
Cambridge, UK
sjc53@cam.ac.uk

Kate Coughlan

BBC
London, UK
kate.coughlan@bbc.co.uk

Kanta Dihal

Leverhulme Centre for the Future
of Intelligence
University of Cambridge
Cambridge, UK
ksd38@cam.ac.uk

ABSTRACT

How AI is perceived by the public can have significant impact on how it is developed, deployed and regulated. Some commentators argue that perceptions are currently distorted or extreme. This paper discusses the results of a nationally representative survey of the UK population on their perceptions of AI. The survey solicited responses to eight common narratives about AI (four optimistic, four pessimistic), plus views on what AI is, how likely it is to impact in respondents' lifetimes, and whether they can influence it. 42% of respondents offered a plausible definition of AI, while 25% thought it meant robots. Of the narratives presented, those associated with automation were best known, followed by the idea that AI would become more powerful than humans. Overall results showed that the most common visions of the impact of AI elicit significant anxiety. Only two of the eight narratives elicited more excitement than concern (AI making life easier, and extending life). Respondents felt they had no control over AI's development, citing the power of corporations or government, or versions of technological determinism. Negotiating the deployment of AI will require contending with these anxieties.

CCS CONCEPTS

• Surveys and overviews • Artificial intelligence • Automation

KEYWORDS

Science communication; AI ethics; AI policy; public perception

ACM Reference format:

Stephen Cave, Kate Coughlan and Kanta Dihal. 2019. “Scary Robots”: Examining Public Responses to AI. In *AIES '19: AAAI/ACM Conference on AI, Ethics, and Society Proceedings. AIES'19, January 27–28, 2019, Honolulu, HI, USA*. ACM, New York, NY, USA, 7 pages. <https://doi.org/10.1145/3306618.3314232>



This work is licensed under a Creative Commons Attribution-NoDeriv International 4.0 License.

AIES '19, January 27–28, 2019, Honolulu, HI, USA
© 2019 Copyright is held by the owner/author(s).
ACM ISBN 978-1-4503-6324-2/19/01.
<https://doi.org/10.1145/3306618.3314232>

1 Introduction

How developments in artificial intelligence are communicated to, and perceived by, the general public will critically influence the adoption and use of this technology [5]. It should therefore be a matter of concern that there is evidence to suggest that it is going wrong. In a recent report, the UK House of Lords argues that currently “many of the hopes and the fears presently associated with AI are out of kilter with reality” [19]. This view is shared by some foremost AI researchers, for example, Professor Margaret Boden: “AI’s future has been hyped since its inception. Overly enthusiastic predictions from (some) AI professionals have excited, and sometimes terrified, journalists and commentators” [4].

Both excessively hopeful and excessively frightening narratives can have significant negative societal impacts. Exaggerated expectations for what AI can achieve, and when, risk undermining further research and investment. Misplaced trust in AI technologies has already exposed people to a range of risks, including manipulation, privacy violation, and loss of autonomy [9]. Exaggerated fears, on the other hand, may lead to beneficial systems, such as better medical diagnoses, not being adopted. Both misplaced hopes or fears could lead to misguided regulation: for example, lack of regulation could encourage irresponsible use of AI; stifling overregulation could prevent the development and deployment of applications that would enhance the public good [11]. The public perception of AI is therefore an important ethical issue [15].

Existing research has investigated levels of public awareness and understanding [1,23], or public perceptions [11] and views on risks and benefits of particular applications [3,14]. A third focus of research is public opinion on how these technologies should be shaped in future [10,20]. Some studies have represented a global perspective [2] and included several of the areas mentioned above. Building on this work, we set out to explore directly what the UK public believe AI is, and how much they subscribe to the kind of utopian and dystopian narratives mentioned above.

In what follows, we explain our methodology in section 2 and the results in section 3; discussion is then in section 4; and recommendations for next steps are in section 5, followed by a brief conclusion.

2 Methodology

2.1 Sample

To test the awareness of such narratives among the UK general population, and the likelihood they attribute to these prospects becoming reality, we conducted a quantitative online survey of 1078 respondents. Respondents were members of an online market research panel of over 20,000 members, selected to be representative of the UK population in terms of geography and key demographics, managed by market research agency GfK on behalf of the BBC. All panel members were invited to complete the survey which included questions to identify the composition of the sample. Their responses were then weighted to provide a nationally representative picture of UK society. In analysis, the data was cut by age group, gender and sociodemographic status. The sociodemographic groupings ABC1 and C2DE were used. ABC1 represents consumers in UK society with higher levels of educational qualification and higher income. C2DE represents consumers with lower income and lower levels of educational qualification [7].

Our sample was typical of the UK population in terms of device ownership and attitudes towards technology [17], with 78% owning a smartphone, 70% owning a laptop, 65% owning a tablet, and 42% owning a desktop computer. 1% owned none of the devices listed. Regarding the role that technology played in their lives, 39% felt that the statement “I can’t live without it but sometimes I do need a little break from it” most accurately described them. 1% felt that the statement “technology scares me” best represented them.

2.2 Questions

Respondents were asked a series of multiple choice and open text questions. The questions were presented to each respondent in a set order, one question at a time. Within each multiple-choice question, the range of possible answers was presented to each respondent in a randomized order, to minimize the influence the order could have on the results.

The first two questions were used to gather data on the respondents. The first question asked “Which of the following do you currently own?” and listed nine devices, from smartphone to virtual reality headset, and ‘none of the above’, as possible answers. The second question asked respondents to select ‘which statement best represents the role technology plays in your life?’ The multiple-choice options ranged from ‘I can’t live without it’ to ‘technology scares me’.

The remaining seven questions in the survey focused specifically on AI. Question 3 was “Have you ever heard of Artificial Intelligence (AI)?”. Those who answered ‘no’ or ‘don’t know’ were not asked any further questions. Those who had heard of AI were asked question 4: “How would you describe Artificial Intelligence to a friend?” This was an open text response question, requiring respondents to explain in their own words or to opt out.

Next, respondents were provided with a working definition of AI: “the development of computer systems able to perform tasks normally requiring human intelligence such as visual perception,

speech recognition, decision-making and translation between languages.”

This was followed by several questions where respondents were asked to evaluate eight claims about AI. These claims were drawn from a framework of hopes and fears that underlie much imaginative thinking about intelligent machines, as expounded in a recent article by Cave and Dihal [6]. Their article claims that Anglophone Western narratives about AI fall into four dichotomies that each consist of a hope and a parallel fear.

The four hopes, *Immortality*, *Ease*, *Gratification*, and *Dominance*, are each associated with narratives in which intelligent machines affect society in a transformatively positive manner. Immortality speaks to the basic drive to stay alive and healthy, and refers to how AI might be used in pursuit of this: for example, through personalised medicine and drug discovery. Ease refers to the desire to be free of drudgery, and the hope that AI will increasingly perform many tasks that people do not want to do. Gratification refers to the way one might wish to use that free time -- that is, pursuing whatever constitutes pleasurable activity -- and the role AI could play in fulfilling these desires. Finally, Dominance, or power over others, can be seen as the means to protect this blissful existence through AI contributing to powerful new means of defence and security.

With each of these hopes, Cave and Dihal pair a parallel fear. Thus, the hope for Immortality contains the threat of Inhumanity: in the pursuit of an ever longer lifespan, a person risks losing their identity, becoming more machine than human. Ease threatens to become Obsolescence, as the desire to be free from work becomes the fear of being put out of work. Gratification carries the risk of Alienation when in their desire for artificially perfect interactions, humans become alienated from each other and prefer to interact with machines. And the pursuit of Dominance evokes fears of an Uprising, as a people’s own AI-enabled power turns on them.

All participants were shown all 8 narratives, presented as independent scenarios and in a randomized order, to reduce any possible influence they may have on one another and on our results overall. The positive or negative was implicit in the scenario and left to the respondent to interpret. Question 5 asked whether respondents had heard of these narratives in the media or in conversation, and question 6 asked whether they felt concerned or excited by them. Question 7 asked whether they felt these statements were likely to come true, and question 8 asked whether respondents believed they would feel the impact of the narratives personally within their lifetime. Respondents who felt that none of the narratives would impact them within their lifetime were asked to specify why not in their own words.

The following definitions were used to describe the eight hopes and fears:

Table 1: The eight hopes and fears.

Immortality	AI might revolutionise medicine, treatment and drugs so that we could live forever.
Inhumanity	AI might enhance our bodies so much that we become more machine than human.
Ease	AI might make our day-to-day lives easier because we could ask computers to do more tasks for us.
Obsolescence	AI might mean we become over reliant on machines and replace the need for humans in jobs, relationships and socialising.
Gratification	AI might become the perfect friend, there to listen whenever we need and ready to meet our every desire.
Alienation	AI might cater to all our desires so well that we prefer AI interaction to human interaction.
Dominance	AI might help strengthen our military power because it could provide smarter weapons.
Uprising	AI might enable computers to become more powerful than us.

Finally, all respondents were asked question 9: “To what extent do you agree or disagree with the following statement: ‘I feel I am able to influence how Artificial Intelligence (AI) develops in the future’” using a scale from strongly agree to strongly disagree. Those who selected “disagree” or “strongly disagree” were asked a final question: to explain in their own words why they felt this way, or to choose “don’t know”.

2.3 Limitations

1078 respondents completed the survey. We were not always able to break down results by age or socioeconomic groups due to small sample sizes. Where we could, any notable variances have been stated in the sections below. Also, due to an oversight, there was some conflation of the definitions used for Obsolescence (“AI might mean we become over reliant on machines and replace the need for humans in jobs, relationships and socialising”) and Alienation (“AI might cater to all our desires so well that we prefer AI interaction to human interaction”): Obsolescence was intended to refer primarily to the workplace (jobs) and Alienation to interpersonal interactions (relationships and socialising). This might have obfuscated responses to these two.

3. Results

Results described are those with a p-value of 0.05 or higher. P-value is used to demonstrate that a claim has a high probability of validity, rather than being the result of randomly occurring noise in the data. We use 0.05 as a threshold for the p-value indicating a 95% confidence in validity, or put another way, a less than 1 in 20 chance of being ‘noise’ [21].

3.1 Awareness of AI

85% of respondents claimed to have heard of AI. 11% had not heard of it and 4% answered “don’t know”. Awareness was high across all age groups (varying between 79-90%).

Verbatim analysis showed a range of levels of sophistication in defining AI. In response to the question “How would you describe Artificial Intelligence (AI) to a friend?”, which was answered by 622 respondents, one respondent wrote: “Depends on the friend - I used to be a postdoctoral research fellow in the subject”. 42% of responses referred to computers performing tasks that replicated aspects of human cognition, such as “decision making”, “learning” or “thinking”. 156 respondents, or 25% of responses, mentioned “robots”. 12% of respondents incorporated hopes or fears, such as “Computers doing things instead of people. I hate it,” or “scary robots” as featured in this article’s title.

3.2 Awareness of Positive & Negative Narratives

When asked whether respondents had heard of the narratives in the media or in conversation, all of the narratives described generated some recognition among respondents. The most commonly recognised narratives were Obsolescence and Ease. The least well recognised narrative was Inhumanity with just 13% recognition.

Table 2: Recognition of the hopes and fears narratives.

Narrative	Recognition
Obsolescence	55%
Ease	53%
Uprising	44%
Dominance	30%
Alienation	20%
Immortality	19%
Gratification	16%
Inhumanity	13%
None of the above	6%

As far as possible, we examined the composition of respondents who were aware of each narrative:

Immortality: More respondents aged 55+ were aware of this narrative (23.1%) than younger groups, with only 15.4% of 35-54s aware. The sample for 16-34s was too low to be robust. More men (22.0%) were aware of this narrative than women (14.9%). Differences between socioeconomic groups were not statistically significant.

Ease: Significantly more respondents aged 55+ were aware of this narrative than any other age group (57.3% 55+ vs. 51.1% 16-34 and 49.3% 35-54). Fewer respondents of sociodemographic group C2DE were aware of this narrative (47.9% C2DE vs. 56.2% ABC1). Gender differences were not significant.

Dominance: More men (36.9%) were aware of this narrative than women (22.8%). More respondents aged 55+ were aware (32.6%) than 35-54s (23.0%). Differences between socioeconomic

groups were not significant. The sample of 16-34s for this question was too low to be robust.

Uprising: More men (48.0%) were aware of this narrative than women (39.3%). Fewer C2DEs (38.7%) were aware of this narrative than ABC1s (47.1%). Age group differences were not significant.

The differences in awareness for Inhumanity, Obsolescence, Gratification, and Alienation were not statistically significant.

3.3 Emotional Responses to the Narratives

Respondents were asked to rate how excited or concerned they felt about each narrative on a scale of 1-10, where 1 equalled concerned and 10 equalled excited. Only Ease and Immortality elicited more excitement than concern. Scores 1-3 are counted towards the total percentage 'concerned', scores 8-10 are counted towards 'excited'.

For all six remaining narratives, respondents were more concerned than excited. This therefore includes two narratives categorised above as hopes: Dominance and Gratification. The Obsolescence narrative elicited most concern.

Table 3: Excitement and concern about the narratives.

Narrative	Excitement	Concern	Avg score
Ease	29%	13%	7
Immortality	25%	21%	7
Inhumanity	8%	38%	4
Uprising	7%	45%	3
Alienation	8%	51%	4
Gratification	11%	30%	6
Dominance	17%	34%	5
Obsolescence	8%	51%	3

3.4 Perceived Likelihood of the Narratives

Most respondents felt that four of the narratives were likely to come true: two hopes, Ease and Dominance, and two fears, Obsolescence and Uprising. For the four remaining narratives, more respondents felt that they were more unlikely to come true than likely to come true, although the responses were relatively evenly split. This question was also asked using a 10-point scale, with scores 1-3 counted towards the total percentage 'unlikely to come true', and 8-10 counted towards 'likely to come true'.

Respondents were asked a separate question on whether they believed each narrative would impact them personally within their lifetime. Where narratives were believed to be likely to come true, respondents believed this would happen within their lifetime. There was very little variation between age groups.

Table 4: Respondents' expectations with regard to the likelihood of narratives coming true.

Narrative	Likely	Unlikely
Ease	48%	5%
Dominance	42%	7%
Obsolescence	35%	12%
Uprising	30%	16%
Alienation	18%	25%
Gratification	18%	26%
Immortality	19%	28%
Inhumanity	12%	30%

Table 5. Respondents' expectations with regard to whether the narratives would impact them personally within their lifetime.

Narrative	Impact in my life-time 16-34	Impact in my life-time 35-54	Impact in my life-time 55+
Ease	53%	52%	58%
Dominance	34%	30%	37%
Obsolescence	33%	40%	37%
Uprising	20%	23%	21%
Alienation	12%	15%	10%
Gratification	12%	14%	10%
Immortality	14%	14%	13%
Inhumanity	7%	9%	8%

3.5 Perceived Influence on AI Development

Across all narratives, 61.8% of respondents disagreed that they were able to influence how AI develops in the future. This disempowerment was not related to which narrative(s) the respondents were aware or unaware of. Respondents were asked to explain why they felt unable to influence the development of AI in an open text question; responses can be divided into three categories:

1. Age. Many older respondents expressed the sentiment that their age prevented them both from having their views heard and from being affected by the technology in their lifetime: "Who is going to listen to an 80 year old? !!" One 58-year-old respondent wrote, "AI is being developed for the under-50s". Age was also by far the most answer given (48% of 58 responses) for explaining why none of the eight narratives would apply to them within their lifetime; the oldest respondent providing this answer was 84, the youngest 58, with an average age of 72.

2. Technological determinism. Respondents across all age ranges expressed the idea that the technology is going to develop regardless of attempts to change or inhibit it: "that shit's out of the bottle now". "Advances in technology will continuously happen regardless if it's negative or positive."

3. Not being consulted. 29.6% of respondents stated that their views are neither solicited nor desired: "Who is going to ask me?" These comments often reflected a sense of dissatisfaction

about not being heard more generally: “How does the average person influence the future at all”; “politicians and business never listen”.

Several replies outlined who *is* perceived to be in control:

1. Business. Several responses criticised big business: “companies will go ahead regardless of what the individual thinks.” On the other hand, some respondents saw an opportunity to exert some influence through consumer behaviour: “consumers as a group may have an influence jointly [*sic*], by carefully choosing which AI products they buy”.

2. Research. 30% expressed feeling disenfranchised by virtue of not having the technical expertise to understand and/or develop AI: “People with more brain than [*sic*] I have will do the developing”. There was also a sense of feeling detached from an ivory tower elite: “When have scientists (and computer technicians) ever listened to the public?”

3. Government. There was a frequent expression of distrust of or detachment from politics: “AI will be influenced by Government not man in street”. On the other hand, several respondents pointed out that government influence itself is limited: “Our governments are too weak to reel in the maniacs creating Artificial Intelligence”.

4 Discussion

4.1 Do Non-Experts Have a Distorted Understanding of What AI Is?

As noted above, nearly half of respondents gave a plausible definition of AI: one involving computers (or other artefacts) engaging in cognitive feats (such as thinking or learning). This is beyond what we expected, perhaps reflecting growing awareness of AI in the general population.

But this result is not incompatible with significant numbers of people holding fairly extreme views of what AI could do or cause: first, because it still leaves more than half of respondents with a less accurate picture of what AI is; and second, because broadly accurate ideas of AI as (a variant of) a thinking computer do not preclude holding extreme views about what those thinking machines might do.

Of course, the extent to which any of the eight narratives used in the survey are ‘extreme’ or distortions depends on one’s standpoint: expert views on the potential of AI vary widely [16]. But four of the narratives were couched in terms of thresholds (AI might allow us to live forever, or AI might become the perfect friend) that we would consider extreme. Although these scored lower than the narratives couched in terms of incremental change (AI might make our lives easier or strengthen our military), they all scored above 10% on both recognition and likelihood. 10% to 20% of the population is a large number of people to have such utopian or dystopian views about an increasingly mainstream technology.

Notably, 7.7% of respondents spontaneously expressed anxiety in response to the request to provide an explanation of what AI is. Some elaborated on their fears (e.g., “must admit I find it a bit worrying, feel we might lose control of our lives if AI takes over”), but 36 responses to this question consisted of short and

simple expressions of anxiety, such as “scary” or “creepy”. This goes beyond the data on the perceived likelihood of the negative narratives coming true: those responses suggest that people believe AI *could* go badly, whereas these open responses suggest that a significant minority see AI as *inevitably* or inherently bad.

It is also noteworthy that 25% of respondents explained AI in terms of robots. We consider this a less accurate explanation than those that centred on thinking machines, as robots (i) are often not particularly intelligent, and (ii) are designed to act physically in the world (whereas AI need not). This conflation is understandable, partly because of the significant overlaps between the fields of robotics and AI, and partly because AI is frequently portrayed in film and fiction as embodied [5]. Nonetheless, it could be problematic. Imagining AI as embodied will lend itself to some narratives more than others: it might, for example encourage the public to focus on worries of gun-toting killer robots rather than the real-world challenge of algorithmic bias. Further, the gendering and racializing of humanoid embodied AI can perpetuate stereotypes [18].

Finally, 14 respondents mentioned fiction in their responses. All of the references to specific works were to films and TV series: *AI: Artificial Intelligence* (6), *The Terminator* (2), *Star Trek* (2), *I, Robot* (1), and *Star Wars* (1). This sample size is too small to draw valid conclusions; further work might explore whether such fictional representations are influencing a broader segment of the public, and if so, whether they are doing so in ways that could be considered distorting or extreme.

4.2 Pairs of Hopes and Fears

As noted above, Cave and Dihal have previously posited that these narratives come as dichotomies; that is, pairs of hopes and fears. As also noted above, the survey did not present the narratives in this way, yet these results do provide some evidence for this being perceived by respondents. First, awareness of the hopeful narratives is broadly similar to awareness of the fears with which they are posited to be paired (Obsolescence 55% and Ease 53%; Uprising 44% and Dominance 30%; Alienation 20% and Gratification 16%; Immortality 19% and Inhumanity 13%).

Second, only two of the narratives elicit more excitement than concern (Ease and Immortality), and even in those cases, the proportion of people excited is well below half (29% and 25% respectively). It is particularly notable that two narratives that we have regarded as broadly positive -- Domination and Gratification -- are perceived to be significantly more concerning than exciting. At the same time, some (albeit relatively few) respondents expressed themselves to be *excited* about the *negative* narratives.

These findings -- that awareness of the positive side of a dichotomy is in proportion with the negative side, and that feelings towards both the positive and negative sides are ambivalent -- suggest that, although the narratives as presented in the survey have attempted to pull apart the positive from the negative, there are limits to how much this is possible. This suggests that a significant number of people recognize that there is an underlying scenario common to the positive and negative visions paired in a dichotomy (such as AI taking over more human jobs) and respond according to their predispositions, regardless of how it is framed.

In particular, it is notable that levels of concern are on average higher than levels of excitement, supporting Cave and Dihal's claim that the hopeful narratives "contain inherent instabilities. The conditions required to fulfil each hope also make a dystopian future possible" [6].

It is perhaps unsurprising that the Dominance narrative elicited more concern than excitement. It is easy to imagine that powerful new weapons create anxiety regardless of which country owns them. This might relate to the influence point above (see 4.3): people might not identify with 'their' military, feel they don't have control over their activity, or have ethical concerns about autonomous machines targeting and killing humans.

That Gratification elicits more concern than excitement is more surprising. The definition we used ("AI might become the perfect friend, there to listen whenever we need and ready to meet our every desire") was *prima facie* wholly positive. The negative reaction suggests that respondents are recognising behind the positive framing an underlying scenario which also has strong negative aspects, something suggested by some of the open text responses expressing anxiety about humans being replaced. It would be interesting to test this in other cultures: for example, both in Japan and in South Korea people are much more positive about seeing AI as a friend [5].

4.3 Perceived Influence on AI Development

The findings show clearly that most people do not feel able to influence the future development of AI. In itself, it is perhaps unsurprising that most individuals do not feel they can shape the direction of this large and highly technical industry. However, the verbatim responses to the question asking why they felt this way are revealing. While some cite a lack of technological expertise, others argue that those who are in control of this industry (whether corporations or governments) neither engage with nor care about the views of the ordinary public. This finding is consistent with earlier research that has shown that people would be more likely to support AI applications if they were given more agency [3].

4.4 Differences Between Demographic Groups

The survey revealed differences in awareness of several narratives based on socioeconomic status: ABC1 respondents were significantly more aware of the Ease and Uprising narratives. Limited awareness of the Uprising narratives among respondents with a lower socioeconomic status is surprising, considering the emphasis tabloids place on this narrative, frequently accompanying even rather innocuous AI-related news items with pictures of the T-800 from the Terminator film franchise [5].

There are also significant differences in awareness based on respondents' gender: women were less aware of the Immortality, Dominance, and Uprising narratives. One potential explanation is that these narratives are very common in science fiction, which for the largest part of its history has been aimed explicitly at adolescent male readers and viewers [12,22]

5 Recommendations

The results suggest many further avenues of work, including exploring the sources of these utopian and dystopian perceptions and the impact of alternative framings of AI. We suggest the following themes for further research:

1. Sources of narratives.

A deeper understanding of the hopes and fears for AI could be obtained through qualitative research investigating the sources of these ideas: which narratives have given individual respondents these perceptions? Such research might take the form of the 'What AI Researchers Read' project, with non-experts rather than AI scientists being the subject of investigation [8].

2. Impact of alternative narratives.

At the same time, further research could look into the impact of new, alternative narratives that are less extreme than the ones mentioned in this paper, or that emphasise aspects of AI outside these eight hopes and fears. Further research might test whether the frightening and overpowering image of AI could be mitigated by emphasising real, current applications, or narratives of control and involvement. It is worth investigating whether narratives that make their audiences imagine an active role in the development or deployment of AI can instil this sense of empowerment.

3. Examining the 'influence' question.

While this survey investigated whether people felt able to influence the development of AI in general, further research might investigate whether and how people feel they have control over the role AI plays in their personal lives, and which narratives might impact that. For instance, an important hypothesis to investigate is that the sense of disenfranchisement might stem from the portrayal of AI research. It has been previously shown that fictional narratives about science tend to focus on individual scientists [13]. Similar research focused on news media coverage should investigate the emphasis given to the role of big tech companies, to see how they portray the role of the non-expert.

4. Examining public perceptions of AI across cultures.

This survey has examined the attitudes of a representative section of the UK population. These results should not be extrapolated globally: there is evidence to suggest that attitudes to AI vary between cultures and regions [5]. More research is needed to evaluate public perceptions of AI in other parts of the world. Such research could both explore recognition of these eight narratives (which, as noted above, were distilled from analysis of Anglophone Western portrayals of AI), and develop further frameworks.

6 Conclusion

This paper has explored a nationally representative survey of the UK population on their perceptions of AI, with a particular emphasis on sentiments regarding utopian or dystopian future scenarios. Overall, the findings show that this population have a markedly negative view of this technology: levels of concern were on average higher than levels of excitement across the narratives; concern was higher than excitement even for two 'hopeful' narratives; and 7.7% spontaneously offered negative sentiments instead of explaining what AI is. Negotiating the deployment of AI

will therefore require contending with the fact that in some parts of the world, a majority of people see downsides even in ostensibly utopian portrayals.

ACKNOWLEDGMENTS

The authors would like to thank Lauren Eriser, Jamie Lawrence, Sarah Dillon, Ali Shah and Toshie Takahashi for data analysis, helpful comments, and discussions.

Authors Cave and Dihal were funded by a grant from the Leverhulme Trust via the Leverhulme Centre for the Future of Intelligence.

REFERENCES

- [1] Mimi An. 2017. *Artificial Intelligence is Here - People Just Don't Realize It*. HubSpot.
- [2] ARM and Northstar. 2017. *AI today, AI tomorrow: Awareness, acceptance and anticipation of AI: A global consumer perspective*. Retrieved from <http://pages.arm.com/rs/312-SAX-488/images/arm-ai-survey-report.pdf>
- [3] Brhmie Balaram, Tony Greenham, and Jasmine Leonard. 2018. *Artificial Intelligence: real public engagement*. RSA, London. Retrieved November 5, 2018 from <https://www.thersa.org/discover/publications-and-articles/reports/artificial-intelligence-real-public-engagement>
- [4] Margaret A. Boden. 2016. *AI: its nature and future*. Oxford University Press, Oxford.
- [5] Stephen Cave, Claire Craig, Kanta Dihal, Sarah Dillon, Jessica Montgomery, Beth Singler, and Lindsay Taylor. 2018. *Portrayals and Perceptions of AI and Why They Matter*. The Royal Society, London. Retrieved from <https://royalsociety.org/~media/policy/projects/ai-narratives/AI-narratives-workshop-findings.pdf>
- [6] Stephen Cave and Kanta Dihal. 2019. Hopes and fears for intelligent machines in fiction and reality. *Nature Machine Intelligence* 1, 2 (February 2019), 74–78. DOI:<https://doi.org/10.1038/s42256-019-0020-9>
- [7] Dawn Collis. 2009. *Social Grade: A Classification Tool*. Ipsos MediaCT. Retrieved from https://web.archive.org/web/20160315075958/https://www.ipsos-mori.com/DownloadPublication/1285_MediaCT_thoughtpiece_Social_Grade_July09_V3_WEB.pdf
- [8] Sarah Dillon and Jennifer Schaffer-Goddard. forthcoming. What AI Researchers Read: The Role of Literature in Artificial Intelligence Research. (forthcoming).
- [9] James Dipple-Johnstone. 2018. ICO statement in response to Facebook data breach announcement. *Information Commissioner's Office*. Retrieved November 5, 2018 from <https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2018/09/ico-statement-in-response-to-facebook-data-breach-announcement/>
- [10] Helen Edwards. 2017. There's a fine line between what people want robots to do and not do for them. *Quartz*. Retrieved November 5, 2018 from <https://qz.com/1101460/theres-a-fine-line-between-what-people-want-robots-to-do-and-not-do-for-them/>
- [11] Ethan Fast and Eric Horvitz. 2016. Long-Term Trends in the Public Perception of Artificial Intelligence. *arXiv:1609.04904v2 [cs.CL]* (September 2016). Retrieved August 21, 2018 from <https://arxiv.org/abs/1609.04904>
- [12] Susan Gubar. 1980. C.L. Moore and the Conventions of Women's Science Fiction (C.L. Moore et les conventions de la science-fiction féminine). *Science Fiction Studies* 7, 1 (1980), 16–27.
- [13] Roslynn D. Haynes. 2017. *From madman to crime fighter: the scientist in western culture*. Johns Hopkins University Press, Baltimore.
- [14] IPSOS Mori and Royal Society. 2017. *Public views of Machine Learning: Findings from public research and engagement conducted on behalf of the Royal Society*. IPSOS Mori Social Research Institute, London.
- [15] Deborah G. Johnson and Mario Verdicchio. 2017. Reframing AI Discourse. *Minds and Machines* 27, 4 (December 2017), 575–590. DOI:<https://doi.org/10.1007/s11023-017-9417-6>
- [16] Vincent C. Müller and Nick Bostrom. 2016. Future Progress in Artificial Intelligence: A Survey of Expert Opinion. *Fundamental Issues of Artificial Intelligence* (2016), 555–572. DOI:https://doi.org/10.1007/978-3-319-26485-1_33
- [17] Ofcom. 2017. *Adults' media use and attitudes*. Retrieved from https://www.ofcom.org.uk/_data/assets/pdf_file/0020/102755/adults-media-use-attitudes-2017.pdf
- [18] Jennifer Robertson. 2010. Gendering Humanoid Robots: Robo-Sexism in Japan. *Body & Society* 16, 2 (June 2010), 1–36. DOI:<https://doi.org/10.1177/1357034X10364767>
- [19] Select Committee on Artificial Intelligence. 2018. *AI in the UK: ready, willing, and able?* House of Lords, London. Retrieved from <https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/100.pdf>
- [20] SPACE10. ongoing. Do You Speak Human? Retrieved November 5, 2018 from <http://doyouspeakhuman.com>
- [21] The Pennsylvania State University. 2019. S.3.2 Hypothesis Testing (P-Value Approach). *PennState: Statistics Online Courses*. Retrieved January 16, 2019 from <https://newonlinecourses.science.psu.edu/statprogram/reviews/statistical-concepts/hypothesis-testing/p-value-approach>
- [22] Lisa Tuttle. 2018. Women in SF. *The Encyclopedia of Science Fiction*. Retrieved January 16, 2019 from http://www.sf-encyclopedia.com/entry/women_in_sf
- [23] Weber Shandwick and KRC Research. 2016. *AI-Ready or not: artificial intelligence here we come! What consumers think & what marketers need to know*. Weber Shandwick, New York.