International Journal of Communication 18(2024), 609–633 1932–8036/20240005

Rage Against the Artificial Intelligence? Understanding Contextuality of Algorithm Aversion and Appreciation

TESSA OOMEN♦[1]
JOÃO GONÇALVES
ANOUK MOLS
Erasmus University Rotterdam, The Netherlands

People tend to be hesitant toward algorithmic tools, and this aversion potentially affects how innovations in artificial intelligence (AI) are effectively implemented. Explanatory mechanisms for aversion are based on individual or structural issues but often lack reflection on real-world contexts. Our study addresses this gap through a mixed-method approach, analyzing seven cases of AI deployment and their public reception on social media and in news articles. Using the Contextual Integrity framework, we argue that most often it is not the AI technology that is perceived as problematic, but that processes related to transparency, consent, and lack of influence by individuals raise aversion. Future research into aversion should acknowledge that technologies cannot be extricated from their contexts if they aim to understand public perceptions of AI innovation.

Keywords: artificial intelligence, algorithm aversion, algorithm appreciation, public perceptions, Contextual Integrity, mixed methods

Artificial intelligence (AI) and algorithm-based tools have recently found renewed interest and investment (Burton, Stein, & Jensen, 2020; Russell, 2021). These tools increasingly support decision-making processes (Burton et al., 2020) or other consequential processes, such as facial recognition by law enforcement agencies worldwide (Hamann & Smith, 2019). Renewed interest in AI presents a key moment to analyze the public receptions of this explosive growth. Public perceptions of AI influence adoption and implementation, as well as trust in technologies, organizations, and society (Dietvorst, Simmons, & Massey, 2015; Helberger, Araujo, & De Vreese, 2020; Logg, Minson, & Moore, 2019), making this a salient topic for research.

Previous works on AI development and implementation (e.g., Russell, 2021) often lack empirical investigation of public perceptions. Similarly, studies into perceptions of AI are often based on experiments

Tessa Oomen: oomen@eshcc.eur.nl
João Gonçalves: ferreiragoncalves@eshcc.eur.nl
Anouk Mols: mols@eshcc.eur.nl
Date submitted: 11-19-2022

[1] This study was made possible by funding from the European Union's Horizon 2020 research and innovation program under Grant Agreement No. 101021808.

(e.g., Castelo, Bos, & Lehmann, 2019), limiting external validity. We add to these accounts by framing AI implementations in their broader societal context through the way these were received by news media and audiences on social media. Our preregistered study[2] builds on the concepts of algorithm aversion (Dietvorst et al., 2015) and appreciation (Logg et al., 2019), and explores how contextual features of AI implementation relate to public perceptions, using Nissenbaum's (2009) theorizing of privacy as contextual integrity (CI).

We analyze discussions of seven AI innovations through automated content analysis (ACA) on Facebook and Reddit. In addition, we conduct a thematic analysis of news items from different international news outlets and technical publications, as well as relevant company press releases. We aim to answer the question *"How do contextual features of AI technologies affect algorithm aversion in public responses to AI innovations?"* This enables us to understand how particular sentiments and emotions are related to applications of AI and the contexts in which they are deployed.

Literature Review

AI has long been prominent in public imaginary, with depictions of sentient or apocalyptic machines being traced as early as Homer's Iliad (Cave et al., 2018). Due to exponential increases in computing power and developments in machine learning (ML), many AI applications from science fiction have materialized in everyday life. When dealing with these automated systems, perceptions are relevant to determine the so-called "machine heuristics" (Sundar, 2008, p. 83). These mental shortcuts determine how individuals engage with and evaluate AI. However, the popularity of AI and the diversity of conceptions require a precise definition to avoid ambiguity within this article. Therefore, when discussing AI, this article targets (1) ML algorithms (2) that aim to emulate human tasks (3) and are deployed in nonresearch settings.

Algorithm Aversion and Appreciation

Although long-term analyses indicate positive perceptions of AI (Fast & Horvitz, 2017), ample research indicates that people hesitate toward algorithmic tools (Cabiddu, Moi, Patriotta, & Allen, 2022; Dietvorst et al., 2015). Algorithm aversion was first suggested by Meehl (1954, as cited in Logg et al., 2019), though it previously lacked empirical evidence. Algorithm aversion refers to people's preference for human methods over algorithmic tools, even when the latter are superior (Dietvorst et al., 2015). Dietvorst and colleagues (2015) found that people tend to rely on human forecasters, especially after seeing the algorithm make mistakes: "Errors that we tolerate in humans become less tolerable when machines make them" (p. 2). This disparity might relate to algorithmic trust that, akin to interpersonal trust, may be disproportionally penalized when expectations of algorithmic performance are violated (Prahl & Van Swol, 2017).

Other factors influencing algorithm aversion are beliefs that algorithmic tools are inflexible and potentially unfair because they cannot consider individual circumstances (Castelo et al., 2019; Helberger et al., 2020), or they cannot support subjective tasks (Castelo et al., 2019). In addition, a mix of individual and structural issues come into play, such as incompatible expectations, lack of expertise, inability to

[2] OSF preregistration: https://osf.io/sm739/?view_only=e66c81515f6d41348d397bc7f52c493b

integrate intuitions in decision making (perceived), reduction in the ability to exert individual autonomy on outcomes, lack of incentives to adopt AI, and approaches of decision-making strategies (Burton et al., 2020).

Algorithm aversion is widely supported in academic and popular media, though Logg et al. (2019) claim aversion is too easily accepted as a given. Instead, they argue that algorithm appreciation is a more likely attitude toward AI in everyday life, even suggesting that people tend to appreciate advice from algorithms over other people's judgments. Logg et al. (2019) suggest that algorithm aversion stems from a broader unwillingness to receive advice from others, whether humans or technology, while (overly) trusting their abilities to judge a situation appropriately (Logg et al., 2019).

Algorithm aversion and appreciation connect to algorithmic trust and technology acceptance (Cabiddu et al., 2022). This article does not aim to give a full reflection of these relevant concepts. Important here is that approaches to perceptions on AI tend to focus on either technical characteristics of the AI or on psychological models. However, understanding perceptions of trust might require a better understanding of influences from outside of AI or individual users. Therefore, our study recognizes that "user acceptance of AI is context-dependent" (Molina & Sundar, 2022, p. 2) and examines the contextual features that might affect people's attitudes toward AI. For this purpose, we adopt Nissenbaum's (2009) CI framework.

Contextual Integrity

Nissenbaum's (2009) CI framework was founded in privacy scholarship and was recently applied to computer science research (Benthall, Gürses, & Nissenbaum, 2017). Nissenbaum (2009) focused on appropriate flows of information and indicated that privacy preferences are context specific. The CI framework describes how information exchange among humans, computers, and institutions takes place in particular contexts for which people construct norms of appropriateness based on expectations of how personal information should flow from one party to another. For instance, expectations of how information should be treated in a medical context are characterized by doctor-patient confidentiality, and data processing by smart cars is characterized by drivers providing informed consent to car manufacturers.

Four parameters shape norms of appropriateness: contexts, actors, attributes, and transmission principles. *Contexts* are understood as "structured social settings characterized by canonical activities, roles, relationships, power structures, norms (or rules), and internal values (goals, ends, purposes)" (Nissenbaum, 2009, p. 132), such as workplaces, medical environments, and family homes. Individuals tend to be more averse toward algorithms in risky or high-stakes contexts (Feng & Gao, 2020; Zhang, Pentina, & Fan, 2021). Therefore, we expect that:

H1: Aversion will be lower for low-consequence contexts (entertainment and personal devices) compared with high-consequence contexts (traffic and surveillance).

Actors are senders and recipients of information, such as individuals, companies, or institutions (Nissenbaum, 2009). Earlier research suggested that individuals are less averse to algorithms employed by nonprofits such as governments (Lourenço, Dellaert, & Donkers, 2020). However, this study focused on pension calculations, and it is unclear whether this finding translates to other domains. Therefore, we formulate the following nondirectional hypothesis:

H2: There is a difference in algorithm aversion between commercial and governmental actors.

Attributes refers to the type of information, such as demographics, pictures, or audio recordings (Nissenbaum, 2009). The type of data influences how people evaluate the appropriateness of data flows. Contextual variations in perceptions of different data types connect with research on how people perceive different AI applications. For instance, Mori's concept of the uncanny valley (see Mori, MacDorman, & Kageki, 2012) indicated that people have different perceptions of robots depending on how humanlike they are and how they move. In other words, different representations of artificial humans influence levels of uneasiness (Mori et al., 2012). In different contexts, research on how algorithm aversion depends on the nature of the task showed that individuals are more averse to tasks that are believed to require human intuition and are not objective (Castelo et al., 2019). The cases in this study differ in the data types processed by AI applications, and thus we expect that algorithm aversion varies across data types:

H3: Data attributes will relate to algorithm aversion.

Finally, *transmission principles* constrain the flow of information (such as doctor-patient confidentiality, informed consent, or notice; Nissenbaum, 2009). Informed consent is part of the use of digital and/or AI-driven services whereby users are informed of information flow practices and are given the choice to engage or disengage (Nissenbaum, 2011). Issues of consent have been central to privacy debates and legislation (Solove, 2013). Crucial to consent procedures is that they are trust based. Trust in companies receiving and processing personal information influence user behavior (Acquisti, Brandimarte, & Loewenstein, 2015). Trust in AI owners is connected to trust in appropriate information norms playing out in informed consent and notice procedures. Therefore, we also expect (informed) consent to play a key role in terms of algorithm aversion:

H4: Transmission principles that involve consent will generate less algorithm aversion than those that do not involve consent.

In sum, privacy expectations are embodied in social norms about appropriate information flows determined by the context, involved actors, type of information, and terms and conditions (Nissenbaum, 2009). We borrow the CI framework from privacy scholarship because data processing by AI is also subject to norms of appropriateness. For instance, data sharing between smart cars and law enforcement agencies suggests an inappropriate information flow. Moreover, breaching the norms of appropriateness around data processed by AI can have impacts that go beyond privacy. Whereas smart car companies sharing data with a law enforcement agency can harm the privacy of the persons involved, it can have additional legal repercussions for the driver. Naturally, the concept of appropriateness itself is also context bound and may depend on the expertise and priorities of those who define it. In our study, and in line with previous literature (Lim & O'Connor, 1996), we expect:

H5: Messages on Reddit will be less averse to algorithms than posts on Facebook.

Reddit is generally presented as having a community more passionate about technology than other major social media networks like Facebook (Manicka & Johnson, 2019); thus, this community is less likely to be averse to algorithms.

Finally, Nissenbaum (2019) stated that AI poses new challenges for the CI framework. AI innovations process various data streams that individually mean little but, as an aggregate, might result in potential breaches of norms. For instance, the mouse click that sets a search query about syphilis in motion becomes part of a flow of information guarded by norms of appropriateness (Nissenbaum, 2019). Moreover, AI applications often include data aggregation and derivations that, on itself, adhere to norms of appropriateness but are morally questionable. For instance, when a user consents to a fitness tracker's terms of service, which may involve third-party use, the moral defensibility and user expectations might still be harmed if companies use this data to infer risks of heart disease (Nissenbaum, 2019). By applying the CI framework, we aim to illuminate how the intricacies of AI implementation might evoke algorithm aversion.

Method

Study Design

This study combines quantitative ACA with qualitative in-depth content analysis of seven cases to understand how algorithm aversion differs across contexts. The quantitative exploration focuses on the perception of AI by audiences, including a sentiment analysis and emotion detection analysis of social media posts. The qualitative analysis entails a thematic analysis of news items and case-related press releases from AI owners. This mixed-method approach aims to uncover key aspects of algorithm aversion that were not observable in previous studies.

Case Selection

The seven cases represent exemplary applications of historical technological developments in AI or ML innovation. Case selection occurred based on three criteria: (1) type of ML technology, (2) a critical event brought the case into public discourse, and (3) the case is discussed beyond Northern American or (Western) European contexts.

First, the cases concern at least one of three key streams of ML research and development: computer vision (CV), natural language processing (NLP), or automated decision making (ADM). CV and NLP connect to ML development through the input data type provided (images or text), whereas ADM is application focused and can also integrate CV and NLP in its processes. The three fields are not equivalent or mutually exclusive, but they capture how ML developments are framed both from a public and a developer perspective. For each case, we identified which stream is more prominently featured.

Second, we conducted a Google News search to determine the critical moment, meaning the earliest occasion the case generated news attention. A significant number of news items had to emerge for the case to qualify.

Third, the cases are relevant beyond the context in which they were developed and deployed, and they represent the first or a notable instance where a key ML technology with the potential to be deployed worldwide underwent public scrutiny. For instance, arguments about image rights and usage by Clearview AI recently emerged again with generating algorithms such as DALL-E and Midjourney. The

ChatGPT model was trained by using similar reinforcement learning techniques from AlphaGo. And controversies about fabrication of celebrity statements, as with Bourdain, have resurfaced with an AI-generated interview with former Formula 1 racer Michael Schumacher. These criteria led to the selection presented in Table 1.

Table 1. Case Overview and Details Based on Contextual Integrity Framework.

IBM Watson winning *Jeopardy!* Watson (by IBM) is an AI developed to answer questions posed in natural language on the quiz show *Jeopardy!*, which it won in 2011. Watson encountered some problems with puns or wordplay, but overall it performed well.				
ML Type: ADM	Context: Entertainment (R)	Actor: Commercial (R)	Attributes: Decision (R)	Transmission principles: NA (public data)
AlphaGo winning from Go champion Sedol. In 2016, AlphaGo became the first AI (developed by Google DeepMind) to win a Go competition from a human player. Go is one of the most complex games due to the high number of possible moves.				
ML Type: ADM	Context: Entertainment (R)	Actor: Commercial (R)	Attributes: Decision (R)	Transmission principles: NA (public data)
Tesla's Autopilot first fatal accident. Tesla makes electric cars that have self-driving (adjacent) capabilities. This case is the first fatal accident in May 2016 with a Tesla where autopilot was active.				
ML Type: CV	Context: Traffic	Actor: Commercial (R)	Attributes: Vision	Transmission principles: Consent requested
Apple's Siri privacy breach. In 2019, a whistleblower from Apple disclosed that Siri recordings were reviewed by humans (contractors) and often held confidential data or included sensitive and private moments.				
ML Type: NLP	Context: Personal (R)	Actor: Commercial (R)	Attributes: Audio	Transmission principles: Consent requested
Exposé on Clearview AI. In 2020, the NYT revealed that 600+ law enforcement agencies use Clearview AI, a facial recognition app based on a database with billions of images scraped from social media and other public websites.				
ML Type: CV	Context: Surveillance	Actor: Government	Attributes: Vision	Transmission principles: Consent not requested (R)
Ofqual's algorithm for GCSE grades. In August 2020, GCSE grades were released in the UK. These were not based on exams but were determined by combining teacher judgments and an AI tool. The results negatively affected students who lost their scholarship and/or place at their preferred college.				
ML Type: ADM	Context: Education	Actor: Government	Attributes: Decision (R)	Transmission principles: Consent not requested (R)
Bourdain deepfake in posthumous documentary. The 2021 documentary about Anthony Bourdain contained 50 seconds of AI-generated voice. The voice was used to read out words that Bourdain had written before his death.				
ML Type: NLP	Context: Entertainment (R)	Actor: Commercial (R)	Attributes: Audio	Transmission principles: Consent not requested (disputed; R)

Note. Reference category denoted by (R)

Data Collection

For the automated content analysis, we collected posts from Facebook and Reddit. Facebook posts were collected through CrowdTangle's search functionality, collecting posts from public pages and groups. The search used case-related keywords, including AI name, AI owner, and relevant event. Data collection for each case covered the time between the critical moment (see Table 1) and May 2022. Reddit posts and comments were collected with the PushshiftAPI. Its archive approximates the universe of Reddit to an extent suitable for our purposes and was employed in studies with a similar scope (Savela, Garcia, Pellert, & Oksanen, 2021). Duplicate content was removed after scraping, resulting in 70,759 Facebook posts and 113,878 Reddit posts and comments. Siri was the most frequent case in the sample (91,308), followed by Tesla (54,920), AlphaGo (27,759), Clearview (5,018), Watson in *Jeopardy!* (4,761), Ofqual (525), and Bourdain (346).

Although our data collection is comprehensive for these platforms, they are only partially representative of the broader public. Facebook tends to serve a generally older population (Auxier & Anderson, 2021), whereas Reddit, in a U.S. context, has twice as many men compared with women and caters to highly educated young people. Participation in these platforms is subject to systemic biases that lead to the underrepresentation of certain groups. This impacts the generalizability of our findings, but it also means that in terms of validity, those who participate are also most likely to shape algorithmic development. Thus, the limitations of our sample also reflect the limitations of societal engagement.

For the qualitative study, we collected press releases and news articles using the same set of keywords. We manually collected one or two press releases from each of the AI owners, depending on availability (N = 11). In the Bourdain case, two interviews with the director were treated as press releases. In addition, we collected 10 news articles per case (N = 70), through Nexis Uni, from three types of outlets: (1) international Western technology outlets, (2) international Western news outlets (United States or Europe), and (3) international non-Western outlets (Asian, Middle Eastern, or African). Tech-based publications were collected through the websites of Wired, The Verge, Gizmodo, and TechCrunch. All articles had a publication date on or soon after the critical moment and consisted of a minimum length of 300 words.

Analysis

Algorithm aversion and appreciation were operationalized through sentiment and emotions. Sentiment was measured through the Python TextBlob package, and scores ranged from −1 (extremely negative) to 1 (extremely positive; M = 0.10, SD = 0.20). A BERT-based ML classifier (Devlin, Chang, Lee, & Toutanova, 2019) trained on the GoEmotions dataset (Demszky et al., 2020) facilitated emotion detection. GoEmotion is the most comprehensive publicly available emotions dataset and bears similarity to our sample.

The sample of qualitative data (N = 81) was analyzed using ATLAS.ti, following the principles of thematic analysis. This method focuses on "identifying, analyzing and reporting patterns (themes) within data" (Braun & Clarke, 2006, p. 79) and is flexible if researchers detail their process and choices made during coding (Nowell, Norris, White, & Moules, 2017). We created a list of potential codes based on the theoretical foundations of the study and operationalization for the quantitative study (algorithm aversion, algorithm appreciation, CI, and emotions).

In addition, we conducted open coding to uncover themes that may fall outside of the theoretical frame, as CI has not yet been (extensively) applied to AI. For validity, two team members first double-coded 10 articles of the sample. The resulting lists of codes were compared and, where needed, changed, combined, or split up to improve precision or clarity. Then, codes were redefined where necessary and developed into themes. For example, codes such as *safety*, *"better-than-human" threshold*, social good, and *accuracy* were collected under the theme *Appreciation*. For the theme *Aversion*, example codes are *discomfort*, *function creep*, *imperfect*, and *lack of consent*.

Results

Emotions and Aversion on Social Media

Hypotheses were tested through multilevel linear regression models with random intercepts. Comments and posts (Level 1) were assumed to be clustered within Facebook pages and subreddits (Level 2), ensuring that a single subreddit would not have a disproportionate influence on the results (e.g., a ML subreddit that would be overly optimistic). Variables were dummy coded according to Table 1.

Sentiment averages are presented in Table 2, on a scale from −1 to 1, where higher values indicate more positive sentiment. The data shows that sentiment is slightly positive for all cases, especially so for Siri, Ofqual, and Watson.

Table 2. Average Sentiment per Case (on a Scale Between −1 and 1).

Case	Sentiment
Apple Siri	.13
Ofqual	.12
Watson	.12
AlphaGo	.10
Clearview	.07
Tesla	.06
Bourdain	.01
Total	**.08**

Figures 1 and 2 show the means for emotion detection, per case and platform, for interpretability of the coefficients. When comparing absolute values, however, it should be considered that the likelihood of presence of emotions is also influenced by their frequency in the training dataset, and we therefore encourage interpretation and comparison within emotions and not between emotions (e.g., admiration is highly rated because very frequent in the GoEmotions dataset).

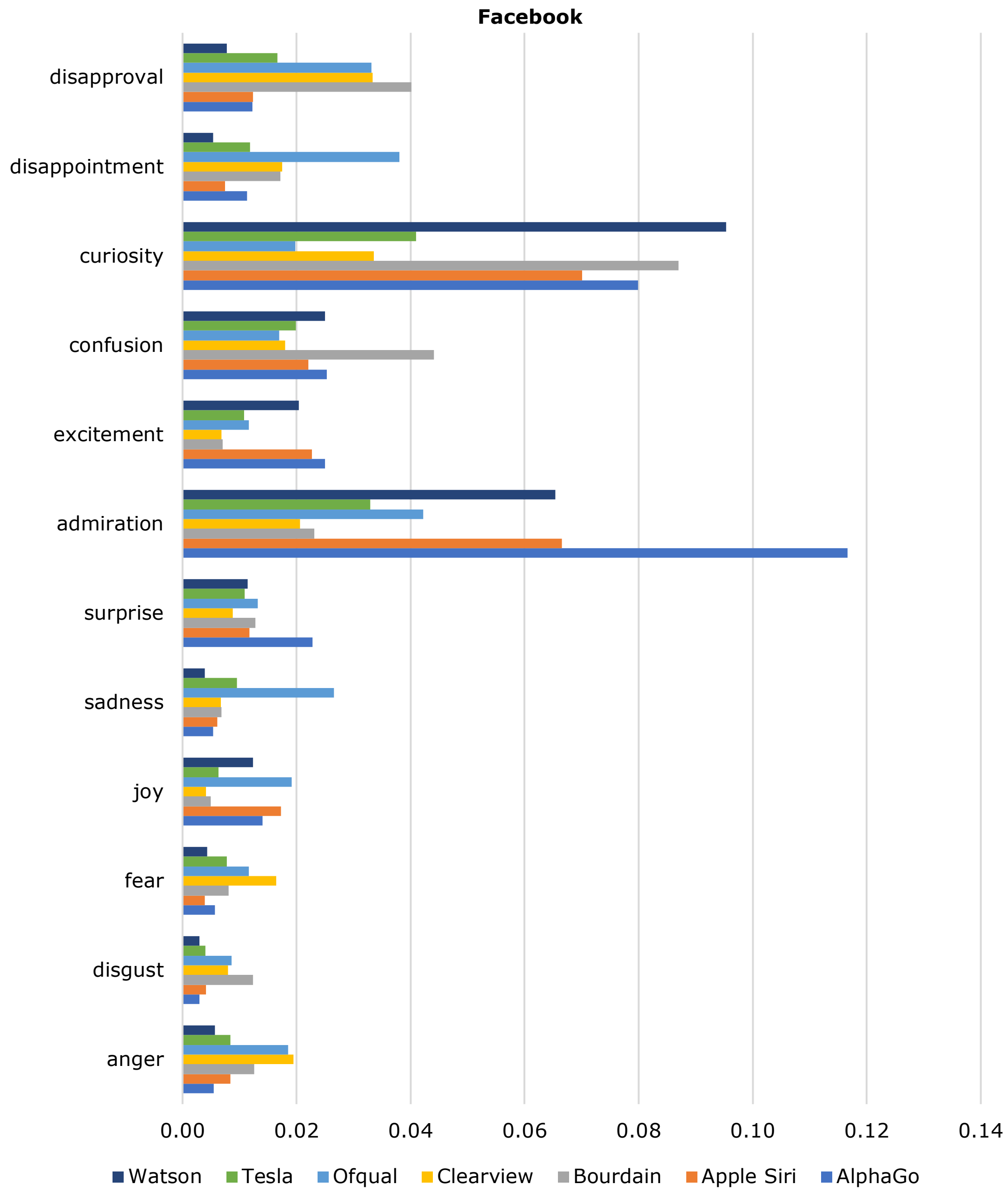

Figure 1. Emotions per case on Facebook.

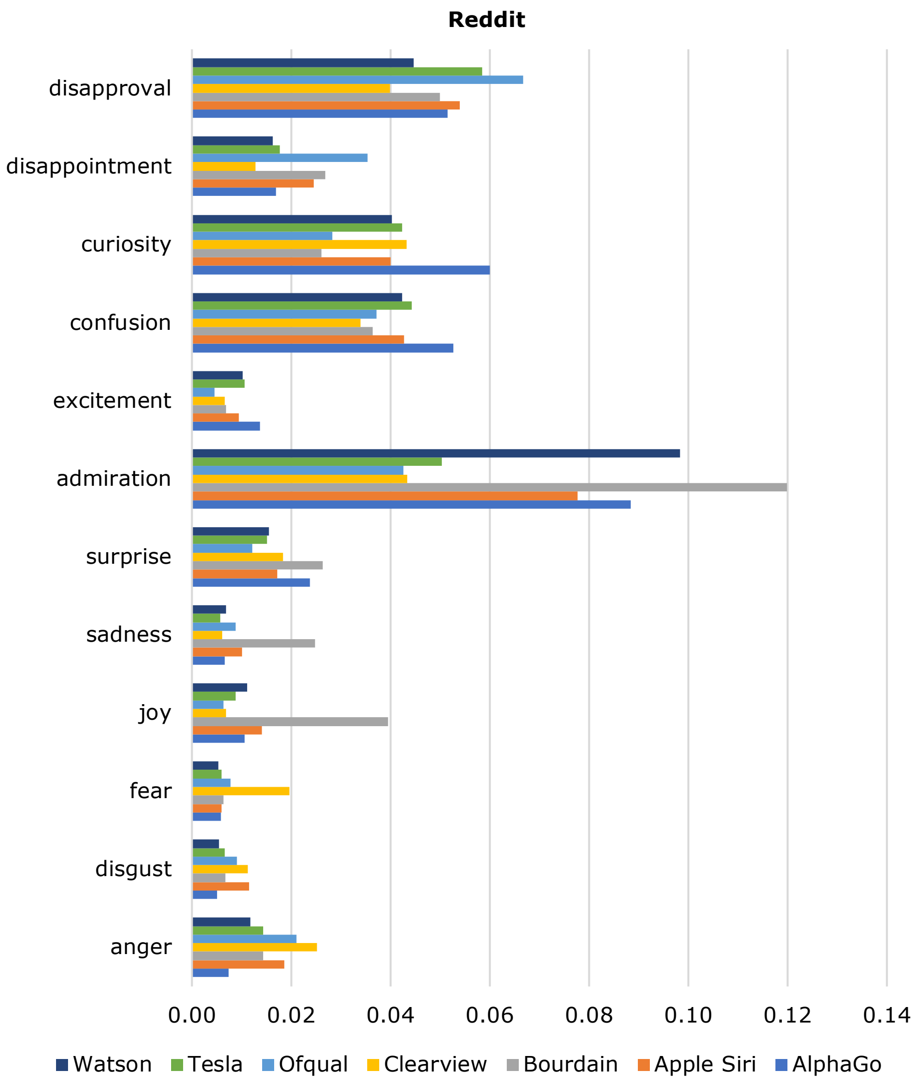

Figure 2. Emotions per case on Reddit.

To test H1, we conducted a multilevel linear regression. The context variable (see Table 1) comprised the independent variable, with sentiment, disapproval, and disappointment as dependent variables. All models are controlled for platform, with Facebook as the reference category. H1 stated that algorithm aversion would be less pronounced for personal devices and entertainment purposes than for traffic and surveillance purposes. This hypothesis is confirmed. Coefficients from Table 3 show that, when compared with personal devices and entertainment, texts about traffic and surveillance applications have significantly (p < .001) more negative sentiment, disapproval, and disappointment.

Table 3. Algorithm Aversion per Context.

	Model 1 **Sentiment**	Model 2 **Disapproval**	Model 3 **Approval**
	Text-Level Variables (Level-1)		
	Coefficient (*SE*)		
Intercept	.10*** (.00)	.01*** (.00)	.05*** (.00)
Education	−01 (.01)	.02*** (.00)	.00 (.01)
Traffic and Surveillance	−.03*** (.00)	.01*** (.00)	−.00*** (.00)
Reddit	.01 (.00)	.04*** (.00)	.05*** (.00)
	Random Effects		
	Variance Component (*SD*)		
Level 2 Variance	0.01 (0.08)	0.00 (0.01)	0.00 (0.01)
AIC (Akaike Information Criterion)	−81,134.31	−387,150.7	−242,929
N of level 1 units (comments/posts)	184,637	184,637	184,637
N of level 2 units (pages/subreddits)	28,537	28,537	28,537

*$p < .05$; **$p < .01$; ***$p < .001$
Note. Reference category: Personal and Entertainment.

Testing H2, a similar process was followed by recoding cases into government or commercial actors. All other model specifications remained the same, as shown in Table 4. The data supports our hypothesis, as discussions of AI usage by government actors show more negativity, disapproval, and disappointment, when compared with commercial actors. Not only is there a significant difference but also all indicators coincide and suggest aversion toward government actors.

Table 4. Algorithm Aversion per Actor.

	Model 4 **Sentiment**	Model 5 **Disapproval**	Model 6 **Approval**
	Text-Level Variables (Level-1)		
	Coefficient (*SE*)		
Intercept	.08*** (.00)	.02*** (.00)	.05*** (.00)
Government	−.03*** (.00)	.01*** (.00)	−.00** (.00)
Reddit	.03*** (.00)	.03*** (.00)	.05*** (.00)
	Random Effects		
	Variance Component (*SD*)		
Level 2 Variance	0.01 (0.08)	0.00 (0.01)	0.00 (0.01)
AIC	−807,32.64	−387,049	−242,915.7
N of level 1 units (comments/posts)	184,637	184,637	184,637
N of level 2 units (pages/subreddits)	28,537	28,537	28,537

*$p < .05$; **$p < .01$; ***$p < .001$
Note. Reference category: Commercial

H3 claimed that data attributes would relate to the perceptions of an ML algorithm. Results in Table 5 support this hypothesis in the sense that video data consistently and significantly shows more aversion than other data types.

Table 5. Algorithm Aversion per Attribute.

	Model 7 **Sentiment**	Model 8 **Disapproval**	Model 9 **Disappointment**
	Text-Level Variables (Level-1)		
	Coefficient (*SE*)		
Intercept	$.10^{***}$ (.00)	$.01^{***}$ (.00)	$.01^{***}$ (.00)
Audio	.00 (.00)	.00 (.00)	.00 (.00)
Visual	$-.03^{***}$ (.00)	$.01^{***}$ (.00)	$.01^{***}$ (.00)
Reddit	$.01^{***}$ (.00)	$.04^{***}$ (.00)	$.04^{***}$ (.00)
	Random Effects		
	Variance Component (*SD*)		
Level 2 Variance	0.01 (0.08)	0.00 (0.01)	0.00 (0.01)
AIC	−81,131.79	−387,126.1	−387,126.1
N of level 1 units (comments/posts)	184,637	184,637	184,637
N of level 2 units (pages/subreddits)	28,537	28,537	28,537

$^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$
Note. Reference category: Decision.

H4 hypothesized a difference in aversion depending on the transmission principles of the ML algorithm. We expected situations where consent was not requested to generate more aversion than those where consent was requested or not applicable. Table 6 shows that, when compared with the no consent reference category, the other types of consent show significantly more positive sentiment, less disapproval, and less disappointment, thus supporting H4.

Finally, H5 posited that texts on Reddit would display less aversion toward algorithms than texts on Facebook. We used multilevel models like the previous hypotheses but included individual cases as a control. H5 is supported for sentiment, showing a positive coefficient for sentiment ($B = .01$, $p < .001$) but not for disapproval ($B = .04$, $p < .001$) and disappointment ($B = .01$, $p < .001$), both of which are higher on Reddit.

Table 6. Algorithm Aversion per Transmission Principle.

	Model 10 **Sentiment**	Model 11 **Disapproval**	Model 12 **Disappointment**
	Text-Level Variables (Level-1)		
	Coefficient (*SE*)		
Intercept	.06*** (.00)	.02*** (.00)	.01*** (.00)
NA	.04*** (.00)	−.01*** (.00)	−.01 (.00)
Requested	.03*** (.00)	−01*** (.00)	−.00*** (.00)
Reddit	.03*** (.00)	.03*** (.00)	.01*** (.00)
	Random Effects		
	Variance Component (*SD*)		
Level 2 Variance	0.01 (0.08)	0.00 (0.01)	0.00 (0.00)
AIC	−80,805.24	−387,090.7	−627,828.8
N of level 1 units (comments/posts)	184,637	184,637	184,637
N of level 2 units (pages/subreddits)	28,537	28,537	28,537

$^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$
Note. Reference category: Not Requested.

Linking Aversion and Contextual Integrity in Cases

The qualitative analysis indicates that news media and AI owners' representations of AI innovations and their public reception include signs of algorithm aversion, appreciation, or both. For each case, we describe the aspects that evoke appreciation and/or aversion and link this to emotions and the CI framework.

Algorithm Appreciation

AlphaGo (2016)

Google DeepMind's AlphaGo works through a neural network using reinforcement learning to compete against itself to defeat professionals at Go. This was not expected for another 10 years (Choe, 2016), which accounts for the admiration and excitement expressed in this case: "[DeepMind researchers] have built a machine capable of something super-human. But at the same time, it's flawed. It can't do everything we humans can do" (Metz, 2016, para. 29). The articles display appreciation of the technology while also emphasizing fact that AlphaGo is fallible.

The attributes of the processed data in this case are publicly accessible expert Go moves and AlphaGo's own moves. The norms of appropriate data processing are based on the premise that AlphaGo only processes expert moves in an entertainment context. However, AlphaGo's AI is already applied for energy efficiency purposes and will potentially support medical research and practice, a contextual change that might affect norms of appropriate data processing.

Watson (2011)

In 2011, IBM's Watson won *Jeopardy!* against two former winners of the American gameshow. Watson processes and responds to questions posed in natural language. The articles showed a sincere appreciation for the technological advances, expressing excitement and admiration. However, most articles were also quick to highlight Watson's limitations and errors, such as regular crashes (Chan, 2011) or seemingly dumb errors.

Watson's nonhuman characteristics were described emphatically ("Man Ties Machine on day 1 of *Jeopardy!* Showdown," 2011; Markoff, 2011). Perhaps that is because natural language is an essential human skill, according to one of Watson's developers (Takahashi, 2011), which could make people question their position in relation to intelligent technologies. In addition, reports featured many jokes and pop culture references highlighting the man-versus-machine debate. These could have been inspired by Jenkins, a contestant, who wrote on the show, "I, for one accept our new overlords" (Markoff, 2011, para. 2), but it might also hint at fear or confusion.

Overall, this case does not include personal or sensitive information, and as a result, the norms of appropriate data processing are not breached. This, combined with the seemingly innocent entertainment context, suggests appreciation in this case. However, the connection to humanlike performance and the jokes may hint at aversion.

Algorithm Appreciation and Aversion

Autopilot (2016)

The case of Tesla shows both appreciation and aversion. The system is commended for its innovation for the automotive industry (Lowy & Seewer, 2016) and improving road safety (Das, 2016). Other indications of appreciation are excitement and joy expressed by Tesla drivers. However, the first (reported) fatal accident with Autopilot has, expectedly, resulted in aversion. At the core, opponents, especially competitors, claim that Autopilot was not sufficiently mature for release nor designed against misuse: "within days [of the launch] people were posting videos of them doing all kinds of stupid things" (Davies, 2016, para. 7). Others express that expecting a perfect system is not realistic (Lowy & Seewer, 2016).

Autopilot applies to traffic contexts, for which legal rules, behavioral norms, and responsibilities apply. Averse responses focused on the lack of responsibility that Tesla showed, stating that Tesla "went out of its way to shift blame for the accident" (Yadron & Tynan, 2016, para. 7) when Tesla stated that Autopilot keeps improving but still requires driver oversight.

Neutral or positive responses discussed the driver's role, because this version of Autopilot "is somewhat like a glorified cruise control" (Das, 2016, para. 3). This comment seemingly downplays the innovation and refers to the responsibility of drivers. For example, the victim was reportedly aware of Autopilot's limitations and might have been distracted at the time of the accident. This would go against Autopilot's user agreement but also the behavioral norms of traffic participation.

In this case, it seems that the technology evokes both appreciation and aversion, but the context and involved actors are raising more aversion.

Algorithm Aversion

Siri (2019)

Apple’s Siri is a virtual assistant that processes voice commands, for which active consent is included through user agreements. Audio processing mainly happens on the device, combined with location data, app data, and contacts, so that Siri can respond more precisely. The norms around appropriate information flows are based on the expectation that users’ data is processed by Siri’s algorithms. These norms were crossed when a whistleblower revealed that Apple employs “human oversight of its automatic voice assistants” (Hern, 2019, para. 20). The transmission principles failed because users were unaware of human actors listening to a pseudonymized selection of audio recordings for improving Siri’s functioning. These contractors listened to “private discussions. Some of these recordings include medical information, drug dealings, and people being intimate” (“Apple QC Workers Often Hear Bits of Private Conversations in Siri Recordings,” 2019, para. 4). The fact that Siri recorded sensitive conversations went beyond user expectations.

At the time of the revelations, “Apple also didn’t offer users any way to opt out beyond disabling Siri altogether” (Byford, 2019, para. 7), hinting at failed transmission principles. By harming the CI of information flows in the use of Siri, Apple faced a transformation from general appreciation of how Siri’s AI functioned to aversion toward how Siri processed audio recordings.

Bourdain’s Deepfake (2021)

The director of Bourdain’s posthumous documentary revealed in an interview that some voice recordings were deepfakes, trained with several audio files that resulted in three clips totaling approximately 50 seconds. Public response was overwhelmingly negative, though some neutral and positive sentiments emerged. Aversion centered around deepfakes’ potentially harmful effects on society, such as diminishing trust and fear of deception (“Posthumous Bourdain Film Raises Questions about Using Data after Death,” 2021), whereas appreciation linked deepfakes to “normal” theatrical trickery in entertainment contexts (Simonite, 2021). Online discussion showed disgust and anger related to how Bourdain would have perceived the deepfakes. Others claimed that Bourdain would have appreciated the unconventionality of deepfakes, indicating neutral and appreciative emotions. In addition, while noting imperfections, audiences admired the quality of the deepfakes.

Ethics were the main source for aversion: “whether it’s ethical to clone a dead person’s voice and have them say things they hadn’t gotten on tape when they were alive is another question” (Gershgorn, 2021, para. 8). The transmission principles would require permission from the subject, which transferred to Bourdain’s estate after his death.

The director claimed to have obtained the estate's permission (Rosner, 2021). Yet audiences questioned the ethicality of recreating the voice of someone who cannot consent (Tangcay, 2021). Relatedly, the audience was not informed about the deepfakes, raising discomfort because they could not make an informed decision about whether they would want to interact with deepfaked media. As per O'Brien and Ortutay (2021), "obtaining consent and disclosing the technowizardry at work would have been appropriate" (para. 8). But the director later stated that "we can have a documentary-ethics panel about it later" (Rosner, 2021, para. 12), which evoked negative sentiments. This case is clustered under aversion because of a breach of the transmission principles and norms of appropriate data processing.

Ofqual's GCSE Algorithm (2020)

The algorithm used to determine final grades for students in the United Kingdom received an overwhelmingly negative public response. Articles reported emotions of anger, sadness, and fear but also confusion about the impact or disputing outcomes (Adams, 2020). Students were angry at the impact of the algorithm's outcomes on their futures, chanting "fuck the algorithm" (Porter, 2020, para. 1) during protests. Governmental or political actors showed anger and disapproval about the choices made during the development process and lack of external evaluation.

Students criticized the lack of opportunity for showing capabilities on an individual basis, a concern reflected by criticism from educators and educational institutions. Experts disapproved of the design and lack of testing of the algorithm because the attributes include historical data categories that may introduce or strengthen bias (Burgess, 2020). Ofqual agreed that the outcomes were faulty and retracted the decisions while also defending their process and position and claiming that multiple stakeholder meetings and evaluations took place.

When considering CI, it is not clear whether breached transmission principles apply. Ofqual was mandated to create the algorithm by the government, and the office of qualifications is expected to manage exams and qualifications. It seems that a breach of norms of appropriate data processing is most applicable for Ofqual's case. When developing an algorithm with extensive impact, a level of care and sound decision making is expected, especially from governmental actors. In addition, the data selection needed further substantiation and review, suggesting an expansion to the norms of CI.

Clearview AI (2020)

The Clearview AI case displays the strongest aversion. Clearview AI's facial recognition tool is used by many law enforcement agencies worldwide. The app processes (incomplete) pictures or footage of faces and returns publicly available images of the same or similarly looking person with links to the websites of origin. Hill (2020) emphasized that the personal attributes of Clearview AI's input and output "end your ability to walk down the street anonymously" (para. 5).

The aversion to Clearview AI revolved around three concerns. First, Clearview AI scraped billions of images from social media and other websites "without providing notice, obtaining informed written consent or publishing data retention policies" (Coldewey, 2020, para. 6), hinting at faulty transmission principles affecting the privacy of Internet users worldwide.

Second, the far-reaching surveillance possibilities were considered problematic, more so due to the risk of false positives and the context of law enforcement use. Potential public availability of the app deepened these concerns. According to Hill (2020), "Searching someone by face could become as easy as Googling a name . . . Someone walking down the street would be immediately identifiable—and his or her home address would be only a few clicks away" (para. 77). Expanding contexts of use will likely increase aversion toward Clearview AI.

Clearview breached the expectations of Internet users worldwide who expected to have control over their images and information online. Clearview AI raised the fear that "facial recognition inherently undermines freedom by enabling perfect surveillance of everyone, all the time" (Hern, 2020, para. 15), suggesting that faulty transmission principles are cause for algorithm aversion.

Discussion

This study aimed to gain a better understanding of algorithm aversion and appreciation through analyzing seven cases of AI implementation by analyzing social media discussions and news reports. The findings of our mixed-method study suggest that algorithm aversion and appreciation are not linear but highly contingent on contextual features. Moreover, although aversion is more likely than appreciation, our results suggest that this is often connected to processes around AI rather than AI itself.

Sentiment analysis indicates that sentiments toward AI technologies are positive though nearly neutral (.01–0.13), although our qualitative data analysis suggests that sentiments more likely to lean toward the negative. This difference may be accounted to the timeline of data collection and the lack of nuance in sentiment analysis. The quantitative data were collected from the critical moment until May 2022, and qualitative data collection covered the day of the critical moment until two weeks after. The longer timeframe for the quantitative data may have allowed for AI owners to respond to public opinion, to right their wrongs, or for stronger emotions to mellow. For instance, much of the positive sentiment in the Ofqual case can be traced to celebrations that the controversial algorithm was abandoned.

These findings indicate the value of combining quantitative and qualitative methods to study algorithm aversion or appreciation. For example, emotion detection indicates admiration as the strongest emotion. In the qualitative data, this is expressed toward technological advances, even when circumstances or actors were judged more negatively. The CI framework gives context to the dynamics of sentiments and accounts for the variations in causes for aversion, though we suggest that there is not one element that determines how audiences perceive AI. Instead, the CI framework allows for an in-depth analysis and detailed case description, highlighting relevant features and tracing causes for aversion or appreciation.

By applying Nissenbaum's (2009) CI framework, we found that algorithm aversion is more likely than appreciation, and that may connect to perspectives on norms of appropriate data processing, as they relate to the four parameters: contexts, actors, attributes, and transmission principles (Nissenbaum, 2009).

Aversion is more likely when the context could impact human life or humanity. Considering aversion in specific contexts, our findings suggest that personal devices and entertainment contexts result in less aversion than traffic and surveillance purposes (H1). Clearview AI infringes on human rights, Ofqual's algorithm affected students' academic careers, and in the Tesla case, death was the main factor for aversion. Bourdain's deepfake raised questions of what happens to digital identities after death. Watson's win, although generally regarded positively, did touch upon (previously) uniquely human capabilities about language, suggesting that entertainment contexts are not automatically positively judged. Contexts hint at issues beyond algorithmic performance as it relates to (diminished) trust in key institutions and processes of society such as law enforcement or the media.

Another factor that might increase aversion more so than commercial actors is the involvement of governmental actors. Negativity, disapproval, or disappointment are more likely regarding governmental actors (H2), which is supported by our qualitative analysis. Although negative sentiments are also expressed toward Tesla and Apple, aversion is more prevalent in cases of governmental use of AI, like Clearview AI and Ofqual. This might suggest a connection to context because governmental decisions are generally more likely to directly impact human life.

The nature of data attributes raises the potential for aversion when the data approximates human likeness. Video data consistently and significantly generated more aversion than other data types (H3). This aligns with the cases of Siri, Bourdain's deepfake, and Clearview AI, all displaying aversion. It is likely that this aversion stems from the personal nature of the data involved because Tesla's use of video data, for example, does not involve personal recordings in this case. When historical data is important, perspectives vary as well. In Ofqual's case, there was strong aversion toward the algorithm's design and the data because they introduced and strengthened bias. Moreover, the predictive value for individual cases of students was very low. Together, they resulted in unfair outcomes, thus evoking aversion. AlphaGo and Watson also used historical data, but both the data and the AI innovation applied to a specific and low-impact context, which seems to be appreciated. Furthermore, over time, trust in algorithms may also play a role (Cabiddu et al., 2022). Individuals may be used to facing algorithmic opponents in settings such as videogames, but they have not encountered them in classroom or educational settings, thus leading to higher levels of aversion.

Finally, flawed, breached, or absent transmission principles may cause an increase in aversion toward algorithmic tools. Cases in which consent is requested or not applicable show significantly more positive sentiments, including less disapproval and less disappointment. The reverse also applies: cases where consent was problematized showed more aversion (H4). This becomes most apparent in the cases of Clearview AI, Bourdain's, and Ofqual's. Our analysis suggests that this is connected to control over AI innovations. In Ofqual's case, students were not able to sit for exams or to control their outcome in another way. For Bourdain's posthumous deepfake, it is impossible to know whether Bourdain would have consented, and audiences were not informed that parts of the audio were a deepfake, taking away their opportunity to choose to engage with faked content. In Siri's case, users were unaware that people had access to recordings, and Clearview AI collected images without informing their subjects.

Applying the lens of CI on AI innovations shows that contexts play an important role in algorithm aversion, whereby the interplay between different actors, the data attributes, and transmission principles shine light on the particularities of adverse sentiments. Previous studies on algorithm aversion (Jussupow, Benbasat, & Heinzl, 2020) and fairness (Starke, Baleis, Keller, & Marcinkowski, 2022) are characterized by contradicting findings, and context is often signaled as a key explanation yet remains unexplored. Our study advances understanding of algorithm aversion by emphasizing that factors that are often disregarded in experiments (ownership, consent, culture) are highly influential in how individuals judge algorithms in practice. In this sense, we show how algorithm aversion goes beyond the technological workings of AI, and rather than treating contextual factors as confounding or control variables, our findings suggest that these should be placed at the core of research designs.

Limitations and Future Research

Our study does not explicitly conform to the original definition of algorithm aversion as the preference toward human actors even when algorithms perform superiorly (Dietvorst et al., 2015). We do not aim at explicit comparison, nor are such direct comparisons clear enough in real-life settings. However, by examining contextual features that shape discourses around AI, particularly about sentiment and emotion, we speak to the literature on aversion and its apparently contradictory findings. Furthermore, our case selection encompasses instances where AI perform tasks that are traditionally human, which supports the use of aversion as a concept.

We expanded the scope of analysis by including cases and data outside of Western-European and Northern-American spheres. However, we were limited to English-language content and encountered redistributions from Reuters through non-Western outlets. Advancing the scope of the analysis requires future research to incorporate languages other than English. The platforms studied in our article, although widely adopted, are also not fully representative of the diversity of publics. It is therefore possible that research undertaken about other (non-Western) platforms would yield different outcomes. Furthermore, our collection of news articles, although useful for generating qualitative insights into our data, is not large or diverse enough to enable generalizations about journalistic interpretations of these critical AI cases.

Conclusion

The contribution of our work lies in three parts. First, our methodological approach directly addresses the call for algorithm aversion studies in real-world settings (Mahmud, Islam, Ahmed, & Smolander, 2022). Our mixed-method approach allowed for contextualizing real-life concerns, how factors interconnect and might affect perceptions of AI, compared with previous (experimental) studies.

Second, our study adds to the literature on algorithm aversion by identifying mechanisms that might induce or strengthen aversion. Burton and colleagues (2020) explain that individual and structural issues can cause aversion, and in relation to public perceptions, our study shows that the likelihood for aversion rises where processes of transparency, evaluation, and consent are lacking. Therefore, linking the literature on algorithm aversion to calls for explainable and trustworthy AI seems promising.

Third, this study contributes to CI literature by expanding the framework's scope from privacy to data processing by algorithms. Norms of appropriateness go beyond privacy as they also entail societal implications. Because CI was originally developed for individual privacy evaluation, our analysis indicates that societal norms are involved, making it worthwhile to expand the scope of CI to AI data processing and to societal CI.

Our study into contextual features of algorithm aversion indicates that errors by AI are noticed yet easily forgiven, whereas errors or bad choices of actors are heavily scrutinized. AI systems thus need careful consideration of design, transparency, and explainability by actors. Furthermore, perceptions of AI cannot be uncoupled from their application to social contexts and norms, which should inform future research in this direction.

References

Acquisti, A., Brandimarte, L., & Loewenstein, G. (2015). Privacy and human behavior in the age of information. *Science, 347*(6221), 509–514. doi:10.1126/science.aaa1465

Adams, R. (2020, August 16). A-levels: Why has Ofqual suspended its criteria to appeal grades? *The Guardian.* Retrieved from https://www.theguardian.com/education/2020/aug/16/a-levels-why-has-ofqual-suspended-its-criteria-to-appeal-grades

Apple QC workers often hear bits of private conversations in Siri recordings. (2019, July 30). *Indian Technology News*. Retrieved from https://advance-lexis-com.eur.idm.oclc.org/api/document?collection=news&id=urn:contentItem:5WP6-X531-JBYT-H3SV-00000-00&context=1516831

Auxier, B., & Anderson, M. (2021, April 7). *Social media use in 2021*. Pew Research Center. Retrieved from https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/

Benthall, S., Gürses, S., & Nissenbaum, H. (2017). Contextual integrity through the lens of computer science. *Foundations and Trends in Privacy and Security, 2*(1), 1–69. doi:10.1561/3300000016

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. doi:10.1191/1478088706qp063oa

Burgess, M. (2020, August 20). The lessons we all must learn from the A-levels algorithm debacle. *Wired UK.* Retrieved from https://www.wired.co.uk/article/gcse-results-alevels-algorithm-explained

Burton, J. W., Stein, M.-K., & Jensen, T. B. (2020). A systematic review of algorithm aversion in augmented decision making. *Journal of Behavioral Decision Making, 33*(2), 220–239. doi:10.1002/bdm.2155

Byford, S. (2019, August 2). Apple stops letting contractors listen to Siri voice recordings and will offer opt-out later. *The Verge.* Retrieved from https://www.theverge.com/2019/8/2/20751270/apple-stops-contractors-siri-voice-recordings-privacy-opt-out

Cabiddu, F., Moi, L., Patriotta, G., & Allen, D. G. (2022). Why do users trust algorithms? A review and conceptualization of initial trust and trust over time. *European Management Journal, 40*(5), 685–706. doi:10.1016/j.emj.2022.06.001

Castelo, N., Bos, M. W., & Lehmann, D. R. (2019). Task-dependent algorithm aversion. *Journal of Marketing Research, 56*(5), 809–825. doi:10.1177/0022243719851788

Cave, S., Craig, C., Dihal, K., Dillon, S., Montgomery, J., Singler, B., & Taylor, L. (2018). *Portrayals and perceptions of AI and why they matter* [Report]. The Royal Society. Retrieved from https://royalsociety.org/-/media/policy/projects/ai-narratives/AI-narratives-workshop-findings.pdf

Chan, C. (2011, February 16). Watson crashed multiple times on *Jeopardy*, plus other Watson tidbits. *Gizmodo.* Retrieved from https://gizmodo.com/watson-crashed-multiple-times-on-jeopardy-plus-other-w-5762296

Choe, S.-H. (2016, March 15). Google's computer program beats Lee Se-dol in Go tournament. *The New York Times*. Retrieved from https://www.nytimes.com/2016/03/16/world/asia/korea-alphago-vs-lee-sedol-go.html

Coldewey, D. (2020, February 14). Class action suit against Clearview AI cites Illinois law that cost Facebook $550M. *TechCrunch*. Retrieved from https://social.techcrunch.com/2020/02/14/class-action-suit-against-clearview-ai-cites-illinois-law-that-cost-facebook-550m/

Das, S. (2016, July 1). Tesla Model S crashes with Autopilot active, driver dies. *Digit*. Retrieved from https://www.digit.in/features/car-tech/tesla-model-s-crashes-with-autopilot-active-driver-dies-30890.html

Davies, A. (2016, June 30). Tesla's autopilot has had its first deadly crash. *Wired*. Retrieved from https://www.wired.com/2016/06/teslas-autopilot-first-deadly-crash/

Demszky, D., Movshovitz-Attias, D., Ko, J., Cowen, A., Nemade, G., & Ravi, S. (2020). GoEmotions: A dataset of fine-grained emotions. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 4040–4054. doi:10.18653/v1/2020.acl-main.372

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019, May 24). *BERT: Pre-training of deep bidirectional transformers for language understanding*. arXiv. doi:10.48550/arXiv.1810.04805

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General, 144*(1), 114–126. doi:10.1037/xge0000033

Fast, E., & Horvitz, E. (2017). Long-term trends in the public perception of artificial intelligence. *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, 31*(1), 963–969. doi:10.5555/3298239.3298381

Feng, X., & Gao, J. (2020). Is optimal recommendation the best? A laboratory investigation under the newsvendor problem. *Decision Support Systems, 131*, 113251. doi:10.1016/j.dss.2020.113251

Hamann, K., & Smith, R. (2019, May 30). Facial recognition technology: Where will it take us? *PCE Publications*, 9–13. Retrieved from https://pceinc.org/facial-recognition-technology-where-will-it-take-us/

Gershgorn, D. (2021, July 15). New Anthony Bourdain documentary deepfakes his voice. *The Verge*. Retrieved from https://www.theverge.com/2021/7/15/22578707/anthony-bourdain-documentary-deepfake-voice

Helberger, N., Araujo, T., & de Vreese, C. H. (2020). Who is the fairest of them all? Public attitudes and expectations regarding automated decision-making. *Computer Law & Security Review, 39*, 1–16. doi:10.1016/j.clsr.2020.105456

Hern, A. (2019, July 26). Apple contractors "regularly hear confidential details" on Siri recordings. *The Guardian*. Retrieved from https://www.theguardian.com/technology/2019/jul/26/apple-contractors-regularly-hear-confidential-details-on-siri-recordings

Hern, A. (2020, January 24). What is facial recognition—And how do police use it? *The Guardian.* Retrieved from https://www.theguardian.com/technology/2020/jan/24/what-is-facial-recognition-and-how-do-police-use-it

Hill, K. (2020, January 18). The secretive company that might end privacy as we know it. *The New York Times.* Retrieved from https://www.nytimes.com/2020/01/18/technology/clearview-privacy-facial-recognition.html

Jussupow, E., Benbasat, I., & Heinzl, A. (2020). Why are we averse towards algorithms? A comprehensive literature review on algorithm aversion. *Proceedings of the 28th European Conference on Information Systems (ECIS), 168*, 1–16. Retrieved from https://aisel.aisnet.org/ecis2020_rp/168

Lim, J. S., & O'Connor, M. (1996). Judgmental forecasting with interactive forecasting support systems. *Decision Support Systems, 16*(4), 339–357. doi:10.1016/0167-9236(95)00009-7

Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes, 151*, 90–103. doi:10.1016/j.obhdp.2018.12.005

Lourenço, C. J. S., Dellaert, B. G. C., & Donkers, B. (2020). Whose algorithm says so: The relationships between type of firm, perceptions of trust and expertise, and the acceptance of financial robo-advice. *Journal of Interactive Marketing, 49*(1), 107–124. doi:10.1016/j.intmar.2019.10.003

Lowy, J., & Seewer, J. (2016, July 1). Death sparks "Autopilot" car probe; man had speeding tickets. *Associated Press*. Retrieved from https://apnews.com/article/fa00456421e441bdb73054031b54e61b

Mahmud, H., Islam, A. K. M. N., Ahmed, S. I., & Smolander, K. (2022). What influences algorithmic decision-making? A systematic literature review on algorithm aversion. *Technological Forecasting and Social Change, 175*, 121390. doi:10.1016/j.techfore.2021.121390

Manicka, A., & Johnson, A. (2019, January 8). *Technology on Reddit*. Retrieved from https://www.redditinc.com/blog/technology-on-reddit/

Man ties machine on Day 1 of *Jeopardy!* showdown. (2011, February 16). *Shenzhen Daily*. Retrieved from http://www.szdaily.com/content/2011-02/16/content_5345174.htm

Markoff, J. (2011, February 16). Computer wins on *Jeopardy!*: Trivial, it's not. *The New York Times*. Retrieved from https://www.nytimes.com/2011/02/17/science/17jeopardy-watson.html

Metz, C. (2016, March 16). In two moves, AlphaGo and Lee Sedol redefined the future. *Wired*. Retrieved from https://www.wired.com/2016/03/two-moves-alphago-lee-sedol-redefined-future/

Molina, M. D., & Sundar, S. S. (2022). Does distrust in humans predict greater trust in AI? Role of individual differences in user responses to content moderation. *New Media & Society*. Advance online publication. doi:10.1177/14614448221103534

Mori, M., MacDorman, K. F., & Kageki, N. (2012). The uncanny valley [From the field]. *IEEE Robotics & Automation Magazine, 19*(2), 98–100. doi:10.1109/MRA.2012.2192811

Nissenbaum, H. (2009). *Privacy in context: Technology, policy, and the integrity of social life*. Redwood City, CA: Stanford University Press. doi:10.1515/9780804772891

Nissenbaum, H. (2011). A contextual approach to privacy online. *Daedalus, 140*(4), 32–48. doi:10.1162/DAED_a_00113

Nissenbaum, H. (2019). Contextual integrity up and down the data food chain. *Theoretical Inquiries in Law, 20*(1), 221–256. doi:10.1515/til-2019-0008

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods, 16*(1), 1–13. doi:10.1177/1609406917733847

O'Brien, M., & Ortutay, B. (2021, July 16). Why the Anthony Bourdain voice cloning creeps people out. *Associated Press*. Retrieved from https://apnews.com/article/anthony-bourdain-documentary-voice-cloning-technology-1dae37f748a22c946e2193fbb00ccc11

Porter, J. (2020, August 17). UK ditches exam results generated by biased algorithm after student protests. *The Verge*. Retrieved from https://www.theverge.com/2020/8/17/21372045/uk-a-level-results-algorithm-biased-coronavirus-covid-19-pandemic-university-applications

Posthumous Bourdain film raises questions about using data after death. (2021, July 25). *The Straits Times*. Retrieved from https://www.straitstimes.com/life/what-should-happen-to-our-data-when-we-die-bourdain-film-raises-ethical-questions

Prahl, A., & Van Swol, L. (2017). Understanding algorithm aversion: When is advice from automation discounted? *Journal of Forecasting, 36*(6), 691–702. doi:10.1002/for.2464

Rosner, H. (2021, July 15). A haunting new documentary about Anthony Bourdain. *The New Yorker*. Retrieved from https://www.newyorker.com/culture/annals-of-gastronomy/the-haunting-afterlife-of-anthony-bourdain

Russell, S. (2021). The history and future of AI. *Oxford Review of Economic Policy, 37*(3), 509–520. doi:10.1093/oxrep/grab013

Savela, N., Garcia, D., Pellert, M., & Oksanen, A. (2021). Emotional talk about robotic technologies on Reddit: Sentiment analysis of life domains, motives, and temporal themes. *New Media & Society*. Advance online publication. doi:10.1177/14614448211067259

Simonite, T. (2021, August 23). Are these the hidden deepfakes in the Anthony Bourdain movie? *Wired*. Retrieved from https://www.wired.com/story/these-hidden-deepfakes-anthony-bourdain-movie/

Solove, D. J. (2013). Introduction: Privacy self-management and the consent dilemma. *Harvard Business Review, 126*(7), 1880–1903.

Starke, C., Baleis, J., Keller, B., & Marcinkowski, F. (2022). Fairness perceptions of algorithmic decision-making: A systematic review of the empirical literature. *Big Data & Society, 9*(2), 1–16. doi:10.1177/20539517221115189

Sundar, S. S. (2008). The MAIN Model: A heuristic approach to understanding technology effects on credibility. In M. J. Metzger & A. J. Flanagin (Eds.), *Digital media, youth & credibility* (pp. 73–100). Cambridge, MA: The MIT Press.

Takahashi, D. (2011, February 17). IBM researcher explains what Watson gets right and wrong. *VentureBeat*. Retrieved from https://venturebeat.com/business/ibm-researcher-explains-what-watson-gets-right-and-wrong/

Tangcay, J. (2021, July 15). Anthony Bourdain voice re-creation in documentary raises controversy. *Variety*. Retrieved from https://variety.com/2021/artisans/news/anthony-bourdain-fake-voice-roadrunner-documentary-backlash-1235020878/

Yadron, D., & Tynan, D. (2016, June 30). Tesla driver dies in first fatal crash while using autopilot mode. *The Guardian*. Retrieved from https://www.theguardian.com/technology/2016/jun/30/tesla-autopilot-death-self-driving-car-elon-musk

Zhang, L., Pentina, I., & Fan, Y. (2021). Who do you choose? Comparing perceptions of human vs. robo-advisor in the context of financial services. *Journal of Services Marketing, 35*(5), 634–646. doi:10.1108/JSM-05-2020-0162