Automated Emotion Recognition in the Workplace: How Proposed Technologies Reveal Potential Futures of Work

Data Inputs, Promised Outputs, and Proposed Actions in Emotion Recognition Patents

KAREN BOYD, San Diego Regional Policy and Innovation Center, USA,[1] karen@sdrpic.org
NAZANIN ANDALIBI, University of Michigan School of Information, USA, andalibi@umich.edu

Emotion recognition technologies, while critiqued for bias, validity, and privacy invasion, continue to be developed and applied in a range of domains including in high-stakes settings like the workplace. We set out to examine emotion recognition technologies proposed for use in the workplace, describing the input data and training, outputs, and actions that these systems take or prompt. We use these design features to reflect on these technologies' implications using the *ethical speculation* lens. We analyzed patent applications that developed emotion recognition technologies to be used in the workplace (*N*=86). We found that these technologies scope data collection broadly; claim to reveal not only targets' emotional *expressions*, but also their *internal* states; and take or prompt a wide range of actions, many of which impact workers' employment and livelihoods. Technologies described in patent applications frequently violated existing guidelines for ethical automated emotion recognition technology. We demonstrate the utility of using patent applications for ethical speculation. In doing so, we suggest that 1) increasing the visibility of claimed emotional states has the potential to create additional emotional labor for workers (a burden that is disproportionately distributed to low-power and marginalized workers) and contribute to a larger pattern of blurring boundaries between expectations of the workplace and a worker's autonomy, and more broadly to the *data colonialism* regime; 2) Emotion recognition technology's failures can be invisible, may inappropriately influence high-stakes workplace decisions and can exacerbate inequity. We discuss the implications of making emotions and emotional data visible in the workplace and submit for consideration implications for designers of emotion recognition, employers who use them, and policymakers.

CCS CONCEPTS • Human Centered Computing → Collaborative and Social Computing → Collaborative and social computing theory, concepts, and paradigms

KEYWORDS: AI Ethics, Emotion Recognition, Emotion AI, Patents, AI and the Future of Work

ACM Reference format:

Karen Boyd, Nazanin Andalibi. 2023. Automated Emotion Recognition in the Workplace: How Proposed Technologies Reveal Potential Futures of Work. In *Proceedings of the ACM on Human-Computer Interaction,* Vol. , CSCW1, Article #95 (October 2023), 40 pages, https://doi.org/10.1145/3579528

1 INTRODUCTION

"The work quality parameters [employee tiredness, negative attitude to work, stress, anger, disrespect, and so forth] can be collected and stored on a server for further display and analysis by employees or employer. For example, an employer can use this information for workforce optimization, laying-off underperforming employees, and promoting those employees that have a positive attitude towards customers. " – Patent 10496947 [130]

[1] The first author was a postdoctoral scholar at University of Michigan at the time this research was conducted

Workers are humans and humans experience emotions. The quote above provides a glimpse into a possible near-future of work with automated, algorithmic, emotion recognition systems that target workers – technologies that collect data about a target person and claim to algorithmically infer their emotions and more broadly affective states [27].

What are these technologies' implications if implemented? Using an *ethical speculation* approach [53], we analyze patent applications (*N*=86) pertaining to emotion recognition and emotion artificial intelligence (AI) in the workplace. We provide a window into possible futures of work with emotion recognition to inform researchers, designers, regulators, and other actors' decisions about sustaining or disrupting these possible futures by answering the following research questions about workplace emotion recognition patent applications:

What *input* and *training* data do they use? What *outputs* do they claim to infer? What *actions* do they take or attempt to prompt?

What are the implications of these inputs, outputs, and actions? Emotion recognition technology in the workplace is growing, and along with it, concerns about the future of work and the workplace [33, 118]. Accelerated by concerns of worker depression, stress, and productivity due to the global COVID-19 pandemic and shifts to remote work (some predicted to continue 'post'-pandemic [160]) by 2024, 50% of organizations are expected to adopt emotion recognition and more broadly emotion recognition technologies [157]. Experts have critiqued automated emotion recognition for constituting a treat to privacy, safety, and autonomy [27, 110]; for rendering inaccurate and opaque results [16, 31, 70]; for adding friction to accountability [31, 70]; for relying on flawed models and proxy data for emotions [16, 137]; and for rendering sometimes biased outcomes and actions [31, 69, 75]. Despite these concerns, emotion recognition technologies are developed, patented, and used in a range of sectors including automotive and industrial safety, law enforcement, entertainment, education, hiring, and worker surveillance [63, 119].

Indeed, emotions are consequential in the workplace [51], and as we argue so are technologies that (claim to) sense emotions in the workplace. Emotions influence job performance, decision-making, behavior, motivation, creativity, turnover, teamwork, negotiation, and leadership [10, 19, 47, 55, 129, 132]. Even fleeting emotions can be the basis of decisions with long-term impact [9]. Further, many occupations (e.g., airline attendants, nurses, customer service, and sales) implicitly or explicitly require workers to regulate their emotional displays and manage others' emotions [72, 138]. Increasingly, the need for workers to manage their emotions (or "emotional labor" [72]) is a recognized aspect of some work forms. Businesses are implementing emotion recognition tools that claim to detect worker emotions [74] with stated goals of assessing and supporting employee compliance, engagement, productivity, and wellbeing [54], or to help workers manage their emotions [95]. As informative as they might be to managers, emotions are also critical to people's sense of privacy [136]. When people decide how and with whom to share their emotions, they make calculated decisions shaped by privacy preferences and contextual norms [8, 21, 117].

The importance and sensitivity of emotions mean that designers and developers of emotion recognition technology make consequential decisions about which emotions are detected, what data is collected, and how inferences of emotions may be used. If technologies are used as proposed by designers, they can have a significant impact on the lives of the workers who are these technologies' targets[2]. Technology descriptions in patent applications offer rich sources of data that reveal detail about not only the features and functionality of novel technology, but also its designers' stated perspectives and objectives [109]. Patent applications allow us to explore "proposed and planned uses," both of which have the potential to impact policy at the highest levels [106]. We posit that analyzing patents facilitates what Fiesler refers to as "ethical speculation"—imagining potential harms of technology, allowing us to conceive of ways to prevent unwanted consequences of its use, for example by advocating for policy change or altering the design to prevent harm before it happens [52]. Articulating these potential harms is important: in order to design or enforce policies about emotion recognition technology, we need a shared understanding of its potential harms [13].

Proposing and making policies and intervening with advocacy necessarily require speculating about uncertain futures [53]. Examining technologies in detail can afford this kind of speculation about the policy and ethical implications of technologies' future use in high-stakes domains [114]. Because patent examining guidelines used by the U.S. Patent Office include novelty, utility, and non-obviousness, we can be confident that patent applications include the novelty and utility imagined by the patent applicants [158]. Patent applications are also particularly useful to examine technologies when technologies' design, deployment, use, and consequences are hidden, not transparent, and unavailable for the public or researchers alike to interrogate [150] as is the case with emotion recognition technologies. Additionally, analyzing patent applications allows us to explore how assignees state aims to influence the future of technologies [150]; indeed while it is "difficult to discern from a patent application the inventions' actual intended uses… the general trend of patents at any given time can suggest the industry's strategic directions and what developments are likely coming." [150].

This paper describes a landscape of emotion recognition technologies intended to be used in the workplace, the utility that patent applicants attribute to these technologies, and their (mis)alignment with existing ethical guidelines for emotion recognition technology. We found that emotion recognition patent applications scoped data collection broadly and claim to reveal "true" "internal" states, including those that a target intends to conceal. We note that, despite calls to avoid doing so [71], many patent applications treated their technologies' estimates ground truth, expecting them to be used as the basis for actions as consequential as hiring, firing, and calling law enforcement. Most patent applications avoided describing training data in detail or at all, giving themselves broad leeway to change and add training data types.

In describing the features of emotion recognition technologies' patent applications and their implications, we perform ethical speculation, discussing potential implications of the technology described in our corpus for the workplace, especially in what its introduction makes more and less visible and how that exacerbates existing inequities in the workplace. We argue that emotion recognition technology's deployment in the workplace creates additional emotional labor, adds privacy threats, and blurs boundaries between worker autonomy and workplace expectations, reinforcing what scholars have called *data*

[2] We refer to the person being monitored with emotion recognition technology at work as the "target" and the company or person who implements the technology or uses its outputs as the "user."

colonialism[3] [143]. We conclude by reviewing possible implications of these technologies' potential failures' *invisibility* in the workplace. In addition to speculation about the futures of work and the workplace that could be enabled by emotion recognition technology, we hope that our descriptions of patent applications provide a foundation for informed ethical speculation about such futures grounded in the particulars of occupations and industries, in concert with other technology, or complemented by data about real-world use.

Computer-Supported Cooperative Work (CSCW) and Human-Computer Interaction (HCI) scholarship has long been interested in sociotechnical understandings of workplace technologies and their implications [6, 66, 67] and lately especially in technologies that quantify and monitor workers [32, 134]. As an emerging technology in the workplace (and beyond), emotion recognition is situated in this longer history and merits attention due to its potential high stakes for workers' dignity and livelihood. On a higher level, this paper has the potential to inform designers', policymakers', and employers' decisions about technology development and use as well as associated policies. Indeed, at its core, this work responds to recent Federal calls seeking to better understand biometric technologies including emotion recognition. For example, in addition to holding a series of public listening sessions in late 2021, in October 2021 the U.S. Office of Science and Technology Policy (OSTP) requested "input from interested parties on past deployments, proposals, pilots, or trials, and current use of biometric technologies for the purposes of identity verification, identification of individuals, and inference of attributes including individual mental and emotional states" [106]. The purpose of this request for information was to "understand the extent and variety of biometric technologies in past, current, or planned use; the domains in which these technologies are being used; the entities making use of them; current principles, practices, or policies governing their use; and the stakeholders that are, or may be, impacted by their use or regulation" [106]. Our work directly speaks to *proposed* and *planned* uses of technologies providing "inference of attributes including individual mental and emotional states" [106] in the workplace domain and its implications – a topic of immense timely and societal relevance.

2 BACKGROUND

This study is informed by research about emotional and affective labor, workplace surveillance, and emotion recognition technology.

2.1 Emotional Labor

Employers can implicitly or explicitly require workers to manage their emotions as part of their work. This can be described as "emotional labor" [72], which can be thought of as part of the related concept "affective labor," which also includes the expectation that one (e.g., a worker) manages the feelings of others [43]. Emotional labor can be accomplished through "surface acting" or "deep acting" [72]. Surface acting involves a person changing only their emotional expression without trying to change their internal experience, while deep acting involves attempting to change internal states. Emotional labor can reduce job satisfaction, increase stress, and increase burnout [26, 112].

[3] The pattern of high-power institutions (like governments and employers) creating and claiming ownership of data about citizens and workers.

Emotional labor expectations are not evenly distributed across jobs. For example, emotional labor plays a key role in the job performance of service workers who interact with customers [89]. Jobs that have been traditionally (and problematically) considered "women's work" in the U.S. tend to have more emotional labor expectations [68]. Emotional labor is undervalued by employers; it is one of the several ways in which service workers and women are more frequently undercompensated [60, 138]. Some argue that emotional labor expectations should be eliminated because they are undervalued, allow and even expect workers to be disrespected by customers, and constitute interactional, procedural, and distributive injustice [60].

Emotion recognition technologies may expand the scope and stakes of emotional labor expectations by claiming to give supervisors reliable, even real-time insight into the internal states of their employees, even as we find when the workers are alone and not expecting or wanting to be watched. We align with scholars like Andalibi and Buss who argue that critically evaluating automatic emotion recognition technologies *before* they are mainstream and normalized can allow us to avoid its potential negative impacts [7].

2.2 Workplace Surveillance

Workplace surveillance in the U.S. includes everything from normalized features of work, like timecards, to Henry Ford's invasive "sociological department" that monitored employees' home lives [4]. Employers monitor employees to prevent theft and fraud, reward and punish employees based on their productivity, enforce workplace safety and conduct policies, interfere with unionization efforts, punish employees who speak negatively about their jobs or employers publicly, or ensure employees reflect the company's values or brand in their personal lives and social media [4, 12, 154].

Many workers object to workplace surveillance broadly, and that objection may be stronger among people in low-power positions [135], who have also historically been more subjected to surveillance practices [77]. Some workers worry that employers will sell data about employees on the market or share data with law enforcement without giving employees meaningful choice [128]. Indeed, workplace surveillance increases worker stress, decreases job satisfaction, increases social isolation, and fosters the perception among workers that work's quantity is more important than its quality [3].

Workplace surveillance is pervasive, invasive, and poorly regulated in the U.S. Employers in many states can and already do record video and audio of employees in their workplace (including employees who work from home), take screenshots of computers at arbitrary intervals, read email, and track their movements outside of work with employer-provided devices or applications [4]. 45% of employers track employees' every keystroke on company devices [144] and passive sensing at work is expanding [104]. U.S. employee privacy protections are a patchwork across states; the development of new tracking technologies has expanded employers' capacity for granular, 24/7 surveillance without commensurate expansion of privacy laws. At this time, only two states (Delaware and Connecticut) require employers to inform employees of electronic surveillance [4]. That said, informing workers, while necessary, does not mean that workers can opt out without consequences. Even so, it appears that many companies do notify their employees, even framing digital workplace surveillance as a form of care for its employees, for example by helping workers grow, increasing safety, fostering a positive workplace, or offering protection to employees [154].

Tracking and analyzing employees' emotions is sometimes legitimized by making a causal link between employee emotions and competitive advantage through intellectual capital, customer service, organizational reactivity, production, or the ability to attract and retain employees [139]. Workplace emotion monitoring tools currently on the market claim they can reduce absenteeism, improve communication, support team cohesion, raise performance, foster creativity, support decision-making and negotiation, encourage organizational citizenship, support health, promote well-being, boost engagement increase productivity, sales turnover, and creativity (as described in design papers [91, 156] and on service websites [50, 100]). However, these claims need to be verified, and regulators and workers should be aware that companies may use emotion recognition software to monitor other aspects of employees' work and policy compliance [93]. In their study of AI-driven pre-employment assessment websites, Raghavan et al. note the long history of bias in hiring and these services' claim that algorithms can address this bias, arguing that the details of the technology and its use are crucial to the question of whether they help or harm job seekers [114] – inspiring our investigation into workplace emotion recognition technology descriptions and important features (i.e., input, output, actions).

Inferring, measuring, and monitoring workers' emotions constitutes an invasive and scalable type of workplace surveillance. Historical use reveals a willingness among organizations to use technology for invasive surveillance. We can monitor and critique the development of emotion recognition systems and advocate for curtailing their use for invasive surveillance with the potential to inform regulation. This study will report the types of data collected by automatic emotion recognition technologies and the actions that the patent applicants imagine that their systems will facilitate, allowing us to consider the implications of this emerging type of surveillance at work.

2.3 Emotion Recognition

Scholars have long debated conceptualizations of emotions [121, 146]. Theoretical approaches to emotions, which emotion recognition technologies can choose to operationalize, include (1) the evolutionary approach [146] within which Basic Emotion Theory (BET) suggests anger, disgust, fear, happiness, sadness, and surprise are the six basic emotions [46] – most emotion recognition technologies rely on this approach; (2) the appraisal approach [146], which includes theories that associate emotions, like arousal (level of excitement provoked by a stimulus), valence (degree of liking towards a stimulus) and sometimes dominance (degree of control exerted by a stimulus) [148]; and (3) the constructionism approach, which considers emotions from a social-psychological stance and takes into account the context in which emotions surface [146].

The idea that algorithms can detect internal emotional states is highly contested [16, 36, 36, 37, 98, 137]. Classifying human emotion from facial expressions is unreliable [16, 17, 78] and evidence suggests automated systems using facial expressions are biased on the basis of race [116]. Similar biases exist in other data used to detect and predict emotion. Automatic speech analysis is also plagued by racial and gender bias [59, 142]. Research indicates that sentiment in text can vary by culture, language, and demographics like gender [127, 151] and that these variations may compromise the accuracy of sentiment analysis [155]. Prior work has examined social media users' perspectives towards emotion recognition highlighting overall negative attitudes [7, 120]. Recent research has also cast doubt on the validity of the emotion models and data used in automated emotion recognition systems [137]. Others point to the technology's "contested scientific foundations" and potential to encode bias, recommending that companies stop using it and governments ban it until more is known about its impacts [36].

Recent work has highlighted the ethical implications of claiming to read internal emotional states. This work has concluded that automated emotion recognition systems "should not claim to determine one's emotional state from their utterance, facial expression, gait, etc." and highlight the importance of ethical analysis at the earliest stages of planning and design [98]. Hernandez et al. recently proposed 12 guidelines for automated emotion recognition with the goals of responsible communication by the system to users, informed consent for data and decision subjects, systems that are finely calibrated to the context of their deployment, and error handling that mitigates harm [71]. While attending to these issues is important, it is unclear if emotion recognition technologies align with these guidelines; furthermore, these guidelines do not draw from empirical analysis.

Emotion recognition algorithms are not perfect. Accuracy rates have improved but may not be high enough to rely on for high-stakes decisions. A recent study improving emotion recognition accuracy using speech data achieved 72.25%, 85.57%, and 77.02% accuracy over three benchmark datasets [103]. A review of facial *recognition* software (i.e. software that compares images of faces to determine whether those images include the same person [124]) cited accuracy rates for 17 algorithms averaging 82.6%, ranging from 50% to 98% (the 98% accurate algorithm was not suitable for real-time applications) [96]. Facial *emotion* recognition often uses the same data [123]—mapped locations of facial features—as these identification algorithms, so many patents in our corpus claimed to be able to both recognize identity and emotions in facial images. Note that even if the emotion recognition aspect worked with perfect accuracy, assigning those emotions to the wrong employee could still constitute a consequential error. Another recent study classifying emotions based on body movement reports accuracy of 90% during walking, 96% during sitting, and 86.66% "in an action-independent scenario." These works provide insights into not only these systems' error rates, but also how accuracy rates can vary when use scenarios are less constrained [2]. Scholars have encouraged context aware affective computing to address these and other problems with emotion recognition technology [149], but recent research examining emotion recognition technology in use in the workplace has noted that automated emotion recognition's accuracy suffers in a workplace context and that adding context did not sufficiently address the accuracy problem [82].

It is important to consider how these accuracy measures are determined. First, accuracy measures are not determined by testing the algorithm against real world data, but on data that was withheld from the training dataset; real-world accuracy is expected to be lower because of differences between the less constrained real-world setting and the setting in which the training data was collected. Training data may have been captured in a natural setting, elicited by a prompt, or produced by an actor. The emotion category or score that is considered "correct" for the purposes of determining accuracy are labelled by a human. Humans are not perfect at recognizing emotions either: emotion recognition ability varies across people [57, 85], emotional display rules vary across cultures [78], and people vary in their emotional experience and expression [17]. If an algorithm had zero error, it would have learned its imperfect training data perfectly: accuracy measures of algorithms are not measuring the absolute accuracy of emotion recognition algorithms, but rather the difference between algorithmic and human emotion recognition.

The quality of automated emotion estimates has been heavily critiqued because it is crucial to the question of whether automatic emotion recognition can be relied on to make decisions about employment and the conditions of work [137]. However, *even if hypothetically* inferences are one day accurate and not biased—a promise that some experts critique [16, 137]—we ask what the implications of inferring emotions and emotion related phenomena may be. Complicating this question is the long-standing and ongoing disagreement among experts about how to conceptualize emotion in the first place [108, 121] or measure it [16, 113]. Scholars disagree whether emotions can be classified as one of a set of basic emotions [46], a location on a set of two or three dimensions (i.e., positive or negative valence; dominance or submission; and a high or low activation) [148], or from a set of emotion concepts perceived by the person experiencing the emotion [18]. Research in automated emotion recognition has noted the failings of Basic Emotion Theory in affective computing [42] and computing scholars question whether they can or and have begun to ask if they *should* infer human emotions [141].

Although we find the literature and external-facing information about existing technology useful to inform our understanding of the underlying technology and the way systems are sold and consequently used, we turn to patent descriptions for detailed feature descriptions of systems designed to employ emotion recognition technology in the workplace that their designers believe to be useful (as required by the U.S. Patent Office (USPTO) assessment criteria [56]) without marketing narratives.

2.4 Patents

Governments can confer on inventors the sole right to make and sell an invention; this license is called a patent. Patent regimes are legitimated by a claim that they promote innovation by protecting the financial incentive to create something new, but the link between financial incentive and innovation is controversial: it is difficult to validate and some argue that it distributes innovation incentive unjustly [23, 92, 122]. Many things are controversially patentable, like life forms, human DNA, and life-saving medicines [109]. Critics say that the current patent system limits, rather than promotes, innovation and fails to include mechanisms for the people whose lives are affected by patenting decisions to be heard in the process of policy-setting [109, 159].

In patents, applicants disclose details about how technologies work and how the applicants believe that they meet the criteria for patentability in the country where they are applying. The USPTO evaluates patent applications on three criteria: a technology must be novel, useful, and non-obvious [56]. Thus, analyzing patent applications can offer not only a deep understanding of how technologies are imagined to operate, but also a view of what patent applicants perceive to be the utility and novelty of their technology. This allows us to connect details about data collection, storage, and use, which companies often obscure or conceal in marketing materials or user experience, to the benefits claimed. Patent applications, as opposed to granted patents, give us a view into many imagined futures of emotion recognition technology.

Patent applications can help in observing and predicting large trends in technology development [1, 39, 84]. They give us a view of technologies past the user interface into the detailed operations of the software and hardware and allow us to analyze technology that may not yet be on the mainstream market and in use. In justifying their claims that their technology is useful, authors of patent applications disclose how they imagine their technology being deployed by users in context. By analyzing patent applications, rather than granted patents or documentation about technologies we can confirm will be implemented, we do not attempt to predict the future of automated emotion recognition technology, but rather to allow for essential "ethical speculation" [53] by regulators, advocates, academics, and practitioners to be informed by a broad set of early signals of designers' conceptions of the technology and its usefulness contained in patent applications. In this paper, we will use a purposeful corpus of patent applications to understand how a particular type of technology (emotion recognition) is described and proposed in a particular context (the workplace), allowing us to consider its implications and potential futures.

3 METHODS

To gather a comprehensive list of emotion recognition patent applications, we queried InnovationQ Plus in December 2020. We decided to use InnovationQ Plus and narrow our search to U.S. applications after discussions with a patent librarian to render a corpus of high-quality, relevant, comparable patent applications. To identify automated emotion recognition patent applications whose authors recognized their potential for use in the workplace, we took a two-step approach. First, we identified a large corpus of automated emotion recognition patent applications. Then, we identified which of those technologies' descriptions include discussion or examples of their technology being deployed in the workplace.

3.1 Initial Query

We initially attempted to narrow our search to workplace applications using Cooperative Patent Classification (CPC) codes or keywords in our query, but we discovered they are not consistently applied enough for us to use them to cleanly identify workplace technologies. Our librarian colleague noted that the market pressure to make a patent claim as broad as possible likely encourages the use of general CPC codes, reducing the usefulness of CPC or keyword searches. We share this detail here both to illustrate our process and to add to our community's repertoire of using patent data for future analyses.

Instead, we ran a general query, seeking any patent application that emphasized automated emotion recognition. To capture patent applications that fit the "automated" requirement, we limited our search to applications that had at least one CPC code that included G06: "Computing, Calculating, Counting." To capture emotion recognition patent applications, we searched for abstracts including related keywords, listed in **Error! Reference source not found.**.[4]

Table 1: Two-criteria Initial Patent Query

Criterion 1: Automated	Criteria 2: Emotion Recognition
CPC Codes limited to G06*	"emotion detection" OR "emotion recognition" OR "emotion prediction" OR "mood detection" OR "mood recognition" OR "mood prediction" OR "affect recognition" OR "affect detection" OR "affect prediction" OR "mental health" OR "mental illness" OR "digital phenotyping" OR "digital phenotype" OR "emotion" OR "mood"

This query was imperfect: it included several patent applications that, for example, allowed workers to select their emotion (rather than automatically detecting it), or aimed to change users' mood without any attempt to measure it. The query may also have excluded some applications that would have been relevant, for example if it implemented existing emotion recognition technology, keywords about emotion may not have made it into the abstract. However, computational applications for which emotion recognition or prediction was significant enough to include in the abstract comprises a useful dataset to understand applicants' conceptions of utility and novelty, and imagined futures for emotion recognition technology.

The initial query described in **Error! Reference source not found.** rendered 1163 patent applications, 284 of which we excluded (by examining the application text) for failing to automate their detection or inference of emotion. We included applications for technology that detected emotions or that accepted and analyzed self-reported emotion, as long as the analysis of self-reported emotions was performed by the technology itself. We excluded technologies that 1) only *generated* emotion (e.g., synthetic emotions for a robot); or 2) allowed humans to select and send indicators of their emotion without any computational analysis or manipulation because, in those cases, the workers determine how their emotions will be understood by the system and its users. Our initial set of computational emotion recognition patent applications included 879 documents. We then proceeded to identify workplace patent applications.

3.2 Identifying Workplace Patent Applications

For the present analysis, we were particularly interested in technologies that were imagined by their authors to automatically recognize the emotions of people in the workplace, including people at work and prospective employees. First, we read the abstract and any introductory sections (Sections preceding the detailed disclosure, often including "Technical field," "Background," and "Summary," "Summary of the Disclosure," or "Disclosure of the Invention") of the patent description, including patent applications whose introductory sections mentioned workplace applications. Then, we used a set of keywords related to

[4] "Affect" was not included on its own, because it expanded the results set significantly with unrelated results due to its additional meaning ("to have an effect on"). For example, the fifth most relevant result as determined by the InnovationQ algorithm for a search with the term "affect" was titled "Scheduled thermal control system." The abstract describes its invention as an operational schedule that manages "a range of factors. . . which can affect refrigeration management."

domains to search the whole patent text. We generated a list of keywords describing the actors and settings that defined our scope, then read applications in the workplace domain to identify other words that patent applicants used to describe the workplace context. We iteratively developed this list: in the early stages of identifying our workplace corpus, to ensure that early data didn't have inconsistent criteria with data collected later, we reviewed earlier patent applications when we identified a new keyword. The keywords we used to identify the workplace domain were "employee," "employer," "manager," "supervisor," "interview," "business," "organization," "company," and "[space] work" (to exclude "network," which was extremely common, but include "work," "workplace," or "worker").

We took special care with applications related to two domains that we found tended to use domain-specific language to refer to work and workers: vehicle systems and chat systems. Vehicle systems were by far most often used to provide navigation or music recommendation in consumer vehicles, but rarely they, for example, recognized attentiveness in truck drivers. To capture these commercial driving applications, we searched for "commercial," "delivery," "cargo," "freight," "truck," and "bus" and searched for "driver" and "operator" to see if those descriptions signaled whether the imagined use scenarios included people driving for work. Chat systems designed for customer service or sales were very often framed as primarily being about monitoring customer emotion, but they frequently had features that allowed them to monitor workers as well. They also often had a specialized vocabulary to refer to workers. To capture the relevant patent applications in this group, we read the early part of the description to identify the language used for workers (for example, "representative," "agent," "operator," or "CSR") and searched for those terms in the application text. We also looked at each instance of "emotions" or "mood" to catch moments of talking about, for example, capturing the mood of "other participants" on a two-party call.

For applications whose introductory sections were very general and domain keywords rendered no results, we searched for phrases that preface an example that could signal how the authors imagine the system being used (like "e.g.", "for example," "scenario," and "embodiment."). Applications that discussed use in a workplace, including as an example, were included in this project's set.

This process led to our final dataset of 86 patent applications that mentioned use for automatically monitoring, recognizing, or predicting workers' and job candidates' emotions.

3.3 Analysis

We closely read each application in its entirety and used qualitative coding [11] to identify excerpts of interest based on our research questions (See Table 2: Topics). Research questions informed our analysis goal of understanding the breadth of data types (*input*), the types of estimates made by systems (*output*), and the *actions* patent authors imagined the systems or users taking based on system outputs. The first author conducted a first round of coding the patents applications' text. Authors met weekly to discuss these emerging codes, to refine them, and to draw connections between them.

Patent texts can be challenging to read and interpret. The authors have formal education in computer science, including machine learning techniques and artificial intelligence; in addition to this technical background, both hold PhDs in sociotechnical fields. They have extensive experience in qualitative and quantitative methods. Their relevant research to this project include examining AI development practices and AI's social implications in high-stakes contexts. With these backgrounds, the authors exceed the threshold of "[people] of ordinary skill in the art"—the audience to whom patent applications are required to be legible [158].

Table 2: Topics

Topic	Description
Input	What input data does the system base its estimates/inferences on?
Training	What training data and procedures are described?
Output	What does the patent claim to detect or predict as output? (e.g., mood, emotion, sentiment)
Actions	What action does the system precipitate or suggest based on the promised output?

Once we labelled the corpus of patent applications, we used open coding [83] to describe the landscape of these features by developing subcodes in order to understand the features of technologies included in our corpus. For example, from patent text coded as "input data," we developed subcodes to identify technologies that collected text, facial images, non-facial biometrics, speech, and so on. These subcodes and their prevalence are described in Findings (Section **Error! Reference source not found.**).

4 LIMITATIONS

We could have missed technologies that will eventually be deployed in a workplace context because they are framed generally in their patent applications (and therefore didn't match our workplace criterion) or because they may be deployed in another context and found to be useful by employers. Our inclusion criteria's breadth (including brief mentions of use on workers or prospective employees) allowed us to capture a robust dataset to address our research questions about applicants' conceptions of the technology; however, we encourage future work on the implications of *already implemented* emotion recognition technologies in workplaces. Our dataset is limited if there is a pattern of managers co-opting non-workplace technology, for example by requiring workers to use emotion recognition technology designed for personal use and surrender data, a type of function creep [86]. Research on actual implementation is necessary to capture technologies that are deployed differently than designers expect or propose as well as cases where firms adopt technologies whose designers did not imagine their potential for use in the workplace. Such an analysis would allow us to answer other important research questions around the perceived usefulness of emotion recognition technologies for managers and the workers' experiences.

The harm or benefit of a technology depends on how they are used, but patent applications do not offer concrete evidence about how technologies *will* be deployed. Future studies should study automated emotion recognition as it is used in the workplace and consider stakeholders' perspectives. For example, while we speculate that these technologies have the potential to shift emotional labor, interview or ethnographic studies with workers subjected to these technologies would provide needed evidence. Still, patent applications offer a unique data source that allows us to examine how technologies are conceived of, imagined, and justified by their creators and what implications they might have for our sociotechnical futures.

While patent applications are a valuable source of data, they have limitations because they are legal documents; that is, they may not fully capture the ways technologists conceive of technologies they create or how they will be presented to potential customers. For example, patent applicants may write intentionally difficult-to-parse language to try to prevent competitors from understanding how their technology works. Future work could investigate the design, development, and deployment of emotion recognition systems, for example by examining research papers, analyzing marketing material of the final products, or conducting ethnographies within development teams or workplaces using the technology.

5 FINDINGS

Our patent analysis revealed a landscape of emotion recognition systems for use in the workplace. Before addressing our key research question, we provide an overview of where patent applications come from, what technologies they describe, and what occupations they target.

All patent applications were filed with the U.S. Patent Office to protect intellectual property in the U.S., but many were filed by inventors in other countries. Our dataset included applicants from the U.S. (52), Japan (14), Korea (3), Germany (3), India (2), Taiwan (2), The Netherlands (2), India (2), Canada, France, England, Switzerland, and Turkey. implying global interest in this technology and leading countries (see Appendix A.1). The application dates range from 1998 to 2020, demonstrating a marked recent increase (see Figure 1). Like other AI-driven technology, this increase may be a result of increased computing capacity and the availability of openly available training datasets [145].

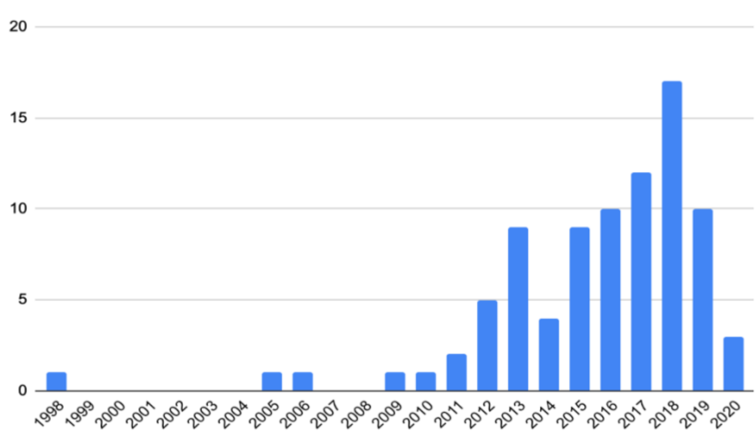

Figure 1: Emotion Recognition in the Workplace Patent Applications by Application Year

While a diverse set of businesses were listed as the assignees of the patents (64 unique entities), a few companies applied for more than one patent in this area, including IBM (16), Samsung (3), Sony, (3), Adobe (2), Hitachi (2), Intel (2), Panasonic (2), SAP (2) Sensory Logic (2), Wipro (2), and 24 7 AI (2), showing a range of large tech companies and smaller startups' stakes and interest in emotion recognition technologies. Three individuals [5] and two Universities are assignees of patents in our corpus as well; see Appendix A.1 for all assignees.

There were several common types of technology described in the patent applications in our set. They most often monitored both customers and company representatives on customer service (and more rarely sales) calls, video calls, chats, and social media exchanges. They also often 1) used sentiment analysis to improve the effectiveness of internal communications (e.g., 0323013, software that suggests tonal changes for work emails); 2) monitored employees' reactions to company policies and events (e.g., 9418390); 3) detected security threats (e.g., 0323013) like workers intending to commit fraud or unauthorized users; 4) detected availability and preferences of workers for meetings or task assignments (e.g., 10169904); and 5) detected workers' feelings in order to change them (e.g., to send targets "cheer up!" messages or attempt to wake a sleepy driver, e.g., 0120219 and 359954).

A plurality of the patent applications did not specify a particular industry or job role (20 applications) or included clauses that expanded their use outside a particular job type (32). We interpret this finding to mean that they are intended to be applicable across jobs and occupations. Customer service was the most common application industry identified (15). Patent applications also often targeted office work (8), social media workers (5), retail (4), drivers (4), healthcare (5) and members of the military (4). Technology was usually designed for use across all work tasks (e.g., monitoring a worker or physical space over time), but some were for use particularly in meetings (7) or during job interviews (11).

In the remainder of this section, we report on results about input, output, and actions proposed in patents.

5.1 Input Data

We found that patent applications described a broad variety of input data and that most applications included more than one potential input data type, but that training data was not well specified.

The technologies in our corpus used a wide range of input data, most commonly text (39 applications), speech (37), facial images (25), non-facial biometrics (24), physical activity (19), video (17), computing behavior of targets (14), and data about context, including sound (8), images (7), video (2), and other context data (11) like weather and light exposure (See Table 3). Most technologies used more than one type of input data. What we refer to here as "other data" varied widely, including the performance of target's favorite sports team, political party, or stocks. Counts of all input data types are reported in Table 3.

Table 3: Input Data Types.

Topic	Description	Amt	Pct
Real-time		147	
Speech	Spoken words, including non-verbal features (e.g., tone, speed)	37	43%
Facial Images	Images of human faces analyzed for their expressions	25	29%
Non-facial biometrics	Any other biological data about the target. (e.g., heart rate, skin temperature)	25	29%

[5] We were not able to identify any research papers by these three individual patent authors. Google Scholar searches for their names and their names along with "emotion recognition" both rendered no results.

Topic	Description	Amt	Pct
Physical activity	Posture, gestures, or body movements	21	24%
Video	Video of people's activity	19	22%
Images	Images of people analyzed for information other than facial expressions, micro-expressions or non-facial biometrics, including images analyzed for information that is not specified	8	9%
Location	The geographic location of the target	7	8%
Eye-tracking	The location and path of the target's eyes	4	5%
Computing Data		71	
Text	Text written by the target, including email and social media data or transcribed speech	40	47%
Computer behavior	Data about how the target uses a computer. e.g., keystrokes, mouse movement	13	15%
Social media	Data about the target's social media	8	9%
Communication Metadata	Data about communication between at least one target and another person. (e.g., frequency of communication, recipients of emails)	8	9%
Mobile data	Computing or unspecified sensor data from a mobile phone	2	2%
Context	Data about the target's environment (excluding people)	32	
Context sound	Ambient sound	9	10%
Context images	Images of a room or other space	8	9%
Context video	Video or a room or other space	2	2%
Personal Data		23	
Context other	Context that is not sound, images, or video (e.g., weather data)	13	15%
Financial data	Information about the target's money	4	5%
Demographics	Information about the target's group membership (e.g., age, gender)	2	2%
Medical data	Information about the target's medical history and diagnoses	2	2%
Family Information	Information about the identities and activities of target's family members	2	2%
Entertainment patterns	The music, movies, television etc. chosen by targets	2	2%
Schedule	Information about the routine of the target or groups that the target is part of	2	2%
Criminal history	Information about a target's criminal record	2	2%
School records	Information about a target's academic history	1	1%
Contact information	Information about the target's social connections	1	1%
Job history	Information about a target's work history	1	1%
Other personal data	Other data about the target that is not otherwise listed	5	6%
Self-Report	Responses to direct questions or unprompted reports about target's emotional state	6	7%
Data about others	Personal or activity data about people other than the target (e.g., customers, co-workers)	2	2%
Sales data	Data about the organizations sales	1	1%
Other	Data not listed in any other category	6	7%

Most (49) technologies described used more than one input type. Some patent applications (18) made long lists of input data types, presumably to avoid limiting the scope of the claimed invention. For example, patent number 0350801 describes using certain data (speech, computing behavior, posture, and biometrics) to detect the "comprehensive state" of workers, interviewees, and customers, and later indicates that the technology can use:

"Social media interactions, facial recognition, global and local events and geopolitical events, financial information, brand affinity, personal preferences, scene analysis, age and gender estimation, professional history, purchase history, navigation traces on web, location history, weather data, event calendar, pre-event and post event status, medical health data, email, subject's family information, subject's psychological information, subject's social connections information, subject's contacts' information, subject's wearable information, subject's physical appearance, subject's crime history, academics data, subject's surroundings information, any other commodities purchased and/or used by the subject, any other information directly or indirectly related to the subject and accessible through the Internet and any other data generated and/or consumed by the subject."

This common approach might be a way for a patent applicant to claim a broad range of appropriate data collection in case they want to use it so that other inventors cannot or to conceal what their real purpose is [150]. However, in practice, this approach leaves space for extremely broad and invasive data collection – in this case among people who may not be in a position to decline data collection.

Patent applications in our corpus made fact claims to support their data collection. For example, Patent 0050837 claims that hemoglobin concentration will reveal the "more than 90%" of emotions that are experienced internally but not revealed through facial expressions. The application does not address whether it is possible to reliably measure hemoglobin concentrations from images of faces in different contexts or whether hemoglobin concentration thoroughly captures internal mental states. Given the potential stakes of these systems, these questions of validity are important to answer before the technologies are patented or put into use.

Many ethical and quality problems in machine learning systems are driven by training data [30] – emotion recognition systems are not an exception. The patent applications in our corpus largely left descriptions of training data very general or avoided them entirely, but a minority of applications described the process of training, specified a training dataset or data source, or specified how training data was labelled. The pattern of either failing to specify or listing a large set of potential training data or features makes it difficult for readers to assess the privacy implications of data collection and the fit of training data to real-world deployments. The lack of specification is not surprising in patent applications, however-- it obscures valuable details from competitors and preserves their options to improve, add, or change training data as development proceeds. It also makes it difficult to truly assess these technologies' "utility." That is, if a technology is invading data collection and use expectations, what does that say about its usefulness? – a criterion patents are evaluated based on. Nevertheless, in this section, we document the patterns of obfuscation and their exceptions. Exceptions highlight ways that training data can be specified to allow readers to evaluate or understand benefits of particular training data without disclosing sensitive information to competitors. We invite reflection on the limitations of assessing technologies' utility or usefulness (a patent assessment criterion [158]) if there is not enough information to assess the validity of proposed technologies.

In our corpus, a plurality of patent applications did not describe their training data at all (41). When applications did discuss training data, they were most often abstract (17), referring for example only to "a historical emotion transcript database" (0358900) or "a 'training set or sample' of user information. . . for which the target property (i.e. an emotional state). . . is known with confidence," (9418390). These patent applications that we considered to have “general” references to training data did not refer to a specific training data set or describe in any detail the features, data source, or labelling of this data.

Some applications introduced more detail about the process of training, but still gave no details or commitment about what data would be used. Patent 9594807 focused on the training process, giving its designers several options for the training process, and left the description of the data itself general: “Non-limiting examples of training techniques that may be used include support vector machine techniques, Fisher discriminant analysis techniques, and neural networks techniques.” It describes the training data only as “facial signatures” that are “provided.” Patent applications frequently described how their models were updated over time (e.g., using personalization (9) or online learning (5)) but did not provide details about how the original model was (or will be) developed (16), limiting the readers’ ability to access the validity and quality of the models.

A few patent applications did offer some detail on training data. Training data sources included data from real-world use (9), best practice datasets (2), prompting emotions in human subjects and measuring their reactions (2), and targets’ web (1) or social media data (1). Annotations came from users of the proposed emotion recognition technology (who were sometimes the targets themselves, and sometimes other parties in the workplace, like interviewers) (7), crowdsourcing (1), or unspecified human annotators (9), while one patent described the annotation as “automated” (1). Patent 0095148 described why it required users (in this case users who are also targets) to annotate their own data, rather than crowdsourcing labels or hiring third party annotators: “Benefits of [user annotation] include the fact that there is no need for manual training, tagging, or manipulation by [people who are searching for emotional content, such as email] The tagging is performed by the author of the textual data used to form the database and train the models, and thus the data derives from organic expression by real users.” Allowing targets to annotate their own data gives targets more control, which could make models, less harmful to targets or more open to manipulation if the targets see the algorithm as a threat.

As another example providing details on training, Patent 0068994 included several examples of training data sources: “For example the information [e.g., labels for training data] could be based on laboratory results, self-reporting trials, and secondary knowledge of emotions (e.g., the individual’s use of emoticons and/or words in their communications). Because some information is more reliable than other information, certain information may be weighted more heavily than other information. For example, in certain embodiments, clinical data is weighted heavier than self-reported data. In other embodiments, self-reported data is weighted heavier than clinical data.” Although this patent is unusually specific about data types compared to the rest of our corpus, it still follows the larger pattern of leaving the assignees a range of options for selecting training data sources and how they are used. Giving examples of training data, however, at least gives us a better picture of how the technology will be built, the implications of using particular sources of training data in proposed use contexts, and the opportunity to critique it from a technical point of view if warranted.

There were a few exceptions to the pattern of avoiding or abstracting training data details. Patent 0116470 used OpenCV as an example of “established methods” their technology would rely on. OpenCV is an open-source computer vision library that offers a training method, through which the inventor provides their own training data, or pre-trained models (the patent does not specify which they will use). Patent 10803255 uses data from real-world use to train the model: "For example, assume a user interaction is manually evaluated [by the user] as a problem . . . Also assume the interaction has a value for each feature to be included in the model (measurements from NLP evaluation, plus anything else system owners have elected to include.) . . . system owners. . .should experiment as they learn more about their problem, and become aware of other indicators to help identify potential problems." This patent includes specific information about the labels used in training data, indicating that they are labeled by system owners as "a problem" or "not", and notes some features' source in the training data (NLP), while preserving the option for system owners to add more features. This quote also includes a human, the system owner, in the loop when updating models. Scholars have criticized human in the loop solutions as a security threat that introduces human failure, a way to ensure that the social realities of sociotechnical systems are taken into account, and a useful construct potentially limited by a lack of domain knowledge, a failure to incorporate more than a limited number of humans' perspectives, and more [29, 35, 115, 131]. In this case, it's worth noting that the "system owner" (likely a worker's manager) is the one labeling interactions as "problematic" and is the one human whose perspective is represented in future system updates, further reifying their power over workers. Here we see that even in where training data details are included, they did not represent approaches that are beneficial to the targets.

Taken together, patent applications in our corpus demonstrated a tendency to scope collection of input and training data broadly and rely on a variety of data types, while not providing meaningful detail about the validity of input and or the nature of training data and processes.

5.1 Output Data

In this section, we describe what applications in our corpus chose to and claim to measure. We find a pattern of surfacing negative or inappropriate emotions, risks of feedback loops, and a tendency to either conflate external states with internal ones or to claim that they are measuring internal states using external signals.

Unsurprisingly given our inclusion criteria, patent applications claimed that their technology would measure emotions (52 patents) as their output, either as binary (e.g., angry or not angry), a location on a two-dimensional scale (activation and valence), from a list of specified emotions (ranging from 3 to 22), or simply "emotions" (of unspecified number and kind). Many measured states of targets (21) (like drowsiness, nervousness, sobriety, conflict, and concentration) or developed scores (17) out of their measurements (for example, Patent 10755712 created scores for voice empathy, keyword empathy, facial empathy, and overall empathy confidence). Patents that measured emotions on a two-dimensional scale relied on the arousal-valence model [148] and those that predicted from a set of specified emotions appear to be inspired by or relying on the same assumptions as Basic Emotion Theory (BET) [108], but more often, they used their own set of emotions, rather than those posed by BET, highlighting inconsistent alignment with emotion theories.

Most patents claimed to measure more than one construct, state, or score. For example, Patent 0228215 measured eight emotion classes as well as attention, engagement, and excitement. Counts for each output type are displayed in Table 4; note that when applications identified affect, we collected the term that the applicants used (e.g. "emotion" or "mood").

Some applications only surfaced negative or “inappropriate” emotions, functionally increasing the visibility of workers’ failure or lack of compliance. For example, 9672825 and 10595764 only detected anger; 0285700 detects conflict; 9105042 tracks negative emotion; and 10496947 detects “employee tiredness, negative attitude to work, stress, anger, disrespect, and so forth.” Although none in our corpus did so, a system could instead focus on surfacing success, and using that information for rewarding workers and instructing trainees. Detecting only negative emotion is a meaningful departure from the emotion theory literature reviewed above which describes a range of emotions, doesn’t prescribe emotions to be rewarded or punished, and often recognizes the contextual nature of human affect [15]. Recognizing the context of emotional expression could matter in the workplace context; for example, if a worker empathizing with a customer’s frustration is read by the technology as expressing a negative emotion and penalized when that expression may in fact be highly-aligned with organizational goals.

5.1.1 Feedback Loops

Many of these technologies are subject to potential feedback loops: in particular, systems that include ongoing training (e.g., personalization or online learning) and that either take their own actions or whose recommendations for human action are consistently acted on. Patent 0036665, which uses emotion data to authenticate social media users (in this context, for business accounts such as social media managers’ accounts, assuming that the tone of one individual’s posts will be similar over time), including marketing or customer service workers, concisely describes this reuse: “If the user’s action is within the threshold, the action [posting to a social media account] is completed . . . and the action is added to the body of knowledge associated with the user and their profile for iterative learning.”

Feedback loops can affect the work lives of targets as well. Some patent applications (e.g., 0160356) use estimated emotions to evaluate and record the employees’ skills to inform how calls are distributed among employees: “a database may include . . . skills data pertaining to a plurality of resources (which may be human agents). . . [and] any other data that may be useful in making routing decisions.” If people demonstrate skill at working with angry customers, they are assigned future calls from angry customers. This aligns well with the presumed interests of the business providing skilled customer service to callers, but does not consider the experience or well-being of a person who has demonstrated this skill or broader implications of this policy on a workforce. At baseline, representatives are sent a mix of angry and not angry customers, but once they demonstrate that they are skilled at dealing with angry callers, they are sent more angry callers. Emotional exhaustion and burnout are problems among call center workers and appear to be exacerbated by the stress and emotional dissonance required to interact with angry customers at work [88, 90]. In this way, feedback loops can compound emotional labor and lead to burnout among those assessed by the technology to be best at their jobs.

Table 4: Claimed Output Categories

Topic	Description	
Identifies Affect	Detects particular affects. In applications: “emotion” (52, or 60%), “mood” (15, 17%), “sentiment” (8, 9%), “tone” (7, 8%), “stress” (4, 5%), group emotions (3, 3%), “feelings” (2, 2%), experienced	97

Topic	Description	
	emotion (2, 2%), “expressed emotion” (2, 2%), “feelings of closeness” (1, 1%), “atmosphere” (1, 1%)	
Describes Affect	Changes in affect (12, 14%), intensity of affect (7, 8%), probability of estimates (5, 6%), component or precursor of a state or affect (3, 3%), what caused a state or affect (1, 1%)	28
State	Identifying states believed to be relevant (27, 31%) and states associated with policies (2, 2%).	29
Score	Rather than detecting a binary affects (e.g. angry or not angry and happy or not happy), assign a score to affects (e.g. anger score of 5 out of 10) or a custom variable of interest (17, 20%)	17
Detects Anomalies	Identifying anomalies (6, 7%), lies (2, 2%), and deviation from routine (1, 1%)	9
Intention	Estimating threat (3, 3%), risk (1, 1%), or other intents (2, 2%)	6
Behavior	Empathy (2, 2%), Share of speech (1, 1%), and pauses in speech (1, 1%)	4
Traits	Personality (5, 6%) or a behavioral profile (1, 1%)	6
Other		16

Although most patent applications assume that the data the technology collects relate directly to and uncomplicatedly reveal a target person’s emotions, patent 10410655 renders probabilities for both “expressed emotions” and “experienced emotions” and explains their reasoning for doing so: “Expressed and experienced emotions may potentially provide information about possible negative consequences in an individual. For example, an experienced feeling of extreme sadness may be expressed as anger. Understanding experienced emotions may also facilitate better communication between an end-user and a human in a conversational user interface. A knowledge of both expressed and experienced emotions may influence some cognitive processes of the end-user and may help the conversation be guided productively.”

5.1.2 Internal and External States

The patent applications that acknowledge the differences between internal and external emotional expression use that acknowledgement to introduce the claim that their technology can detect internal states, unlike humans or existing technology. They claim that this information will help the other party think and reason about the target’s state to try to improve the effectiveness of the conversation.

One patent (0095148) asks targets to label their own data and offers an explanation for including targets themselves in the process:

“the author may be the only individual who truly knows what emotion is actually expressed in the author’s own text. . . In addition, the author is disincentivized to fabricate the textual data, and the tagging process, because the online forum system . . . is designed to encourage networking among users with similar personal stories.”

This explanation reveals a problem with other patent applications that claim to reveal targets’ emotions without allowing the target to see and edit the outputs: experienced emotion can be subtle and individual. However, it also assumes that targets have complete information about their emotions and does not acknowledge that workplace ramifications (e.g., workers may be penalized for expressing or the visibility of some emotions) may change targets' willingness to be open about their emotions.

In summary, with the exception of one patent that asked targets to label their own data, patent applications in our corpus claim to reveal targets' true, internal emotions that a human evaluator could not perceive. This violates guidance [98] against claiming to measure internal states and also fails to respect interests and intentions of targets.

5.2 Actions

Most applicants described some action that their system would generate or prompt human actors to take. We note the consequences of possible bias in the described systems given the actions described. Some of these actions had notably high-stakes potential consequences.

Proposed technologies often generated a visualization or report (36) rendered recommendations to the target (24), expected performance evaluations of workers to be based on the output (19), the system used the emotion estimates to determine the availability of a target for a meeting or to manage their tasks (10), or took actions to change the targets' mood (7) or motivate them (6). Some evaluated or changed a policy or program (5), served targeted advertisements (5), changed the behavior of workforce robots (4), diagnosed or referred targets to clinicians (3) or informed the training of new workers (2).

5.2.1 Consequences

Some actions have what we interpret to be fairly high stakes. For example, 23 patent applications in our corpus discuss actions that could have direct impact on workers' employment, including performance evaluation (19), supervisor interventions (6), invoking security protocols (8), contacting authorities (2), and one patent that directly states "an employer can use this information for workforce optimization, laying-off underperforming employees, and promoting those employees that have a positive attitude towards customers." (10496947). Another patent, 10187254, describes an emotion responsive computer using computing behavior, mobile data ('texts, emails, "tweets," and other electronic messages sent and received by the user's communications device"), medical data ("prescriptions, medical histories, and other medical records"), and data about the users' context, including images, sound, force (presumably force exerted on the device), and temperature. However, in addition to standard mood outputs leading the system to serve tailored content, it can also detect "certain moods indicating emotional concern, such as 'suicidal,' 'riotous,' or 'murderous.'" Based on that estimation, "police, medical personnel, and even therapists may be alerted." This patent uses a broad range of data, including image and audio sensors along with medical data, to make decisions that label targets as potentially dangerous and give that information to authorities. This is a particularly striking potential example of the "black box" at work, as described by Ifeoma Ajunwa, in which, as a condition of employment, workers are required to surrender personal information, submit to algorithmic evaluations that make authoritative claims about them, and be "governed by an invisible, data-created leash," while the employer can avoid transparency about these assessments, as they were made by a computer [5]. When not only the loss of employment, but calling the police is on the line, the ambiguity created by this black box is especially menacing.

Patent 10417484 similarly attempts to detect whether targets have criminal intent and claims to replace expert judgment: “The use of the artificial intelligence and deep learning mechanism eliminates the need for human intervention in determining intent of the subject.” This is part of a pattern of applicants claiming among their technologies’ benefits the ability to avoid human judgment or render real-time decisions (leaving no time for human intervention). This directly contravenes guidance to avoid treating predictions as truth and precludes system design from allowing decision subjects recourse over data collection and predictions [71].

More commonly than estimating criminal intent, patent applications indicated that their output could be used to evaluate targets’ job performance (19). For example, patent number 10496947 used facial images, gestures, speech, and non-facial biometrics to estimate “work quality parameters” such as “employee tiredness, negative attitude to work, stress, anger, disrespect, and so forth.” The authors expect supervisors to use these outputs to make decisions: “An employer can use this information for workforce optimization, laying-off underperforming employees, and promoting those employees that have a positive attitude towards customers.” In some cases, these performance evaluations are expected to be made on very little data, for example relying a single data source (8) (facial expressions (2), text (2), social media data (2), or recordings of phone calls (2)). Making promotion and firing decisions based on thin data and technology that has substantial quality flaws, like inaccuracy, bias, or lack of validity, can compromise effectiveness and lead to people unfairly losing their jobs.

5.2.2 Bias

Bias is a particular threat. If, like we have seen in existing emotion recognition technologies (as reviewed in Background), the systems in this corpus were to perform poorly on a particular worker or group, it could have significant consequences including the following:

- The system could consistently offer workers bad recommendations, like returning inappropriate search results, sending them referrals for inaccurate diagnoses, and pushing frequent and unnecessary calming music, encouraging messages, or inappropriate, contextually-insensitive ads throughout the workday (0087424, 0110727, 0120219, 0187823, 0068994, 0187823, 10187254, 0278413).
- The system could interfere with workers’ work by scheduling inappropriate meetings or tasks, ignoring the workers’ instructions, matching them with more difficult customers, restricting their access privileges or sending them through additional security protocols, prompting unhelpful behavior from workplace robots, or flagging them as ‘risky’ and ‘in need of attention’ (6263326, 0305976, 10237304, 0068994, 0220777, 10803255, 10516701, 10237304, 10452982, 0036665, 10237304, 10542149, 10719796).
- The system could also send frequent or inappropriate notifications to workers’ supervisors, mentors, coworkers, family members, phone contacts, social networks, security, medical professionals, or authorities (8098273, 0350074,10755187,9594807, 10187254, 0285700, 311864, 10803255, 10417484, 10187254).
- The system could penalize workers by subjecting them to additional surveillance, rating their skills consistently lower or different than they are, less frequently rating them highly or recommending rewards, estimating undesirable emotions

in an interview or meeting, like 'defensive' or 'inattentive', reporting that they are lying, rating them as 'untrustworthy' or 'uncreditworthy', or recommending laying them off (10516701, 10481864, 0272976, 0358900, 0210688, 10402918, 9922666, 8098273, 10496947, 0228215, 0050837, 0160959, 0311371, 10496947).

Applicants see their technologies' emotion estimates as the basis for wide range of actions without considering the impact of these actions on the social conditions of work, especially when the estimates they are based on are inaccurate and biased. The impact of these system actions on the conditions of work is significant regardless of who it is aimed at, and if it is disproportionately distributed by skin color, gender, age, disability status, or country of origin for example, it can compound existing systemic injustices. Patent applications consistently defy guidance to allow data subjects recourse and to avoid treating systems' predictions as ground truth on which decisions can be based [71].

6 DISCUSSION

Our findings contribute a landscape of imagined and proposed emotion recognition technologies and how they may be used in the workplace. We describe their input data (along with associated validity concerns and training data and processes), promised outputs, and suggested actions based on outputs. We note that these proposed emotion recognition technologies violate ethical guidelines [71, 98] for the development of such technology, especially encouragements not to treat predictions as truth; not to claim to measure internal states; and to give data and decision subjects opportunities for feedback, customization, and recourse. We also describe how features (like feedback loops) and failures of these proposed technologies could impact workers' experiences and livelihood in the context of workplace power dynamics.

This paper offers a view into the "utility" (i.e., a criterion upon which patents the U.S. Patent Office evaluates patents) that patent applicants attribute to their technologies. Noting that not all technologies described will necessarily be implemented, they nevertheless reveal a view into technology designers' conceptions and justifications about what technologies are useful, novel, non-obvious, and likely marketable. We hope to begin the discussion here about the possible futures of and with automated emotion recognition technology and its implications in the workplace, and we hope our findings will help researchers, policymakers, and advocates to anticipate and intervene in these possible futures.

In this section, we explore possible futures that the technologies described in our corpus could be a part of or precipitate. Ethical speculation allows us to look into several possible futures and consider what interventions could prevent harm [52]. Grounding our ethical speculation in patent applications allows us to be informed by emotion recognition technologies' designers' conceptions of the technology. Within this ethical speculation framework [53], we argue that, if implemented, technologies described in our corpus have the potential to increase the visibility of (claimed) emotional states, thereby creating additional emotional labor expectations for workers, threatening workers' privacy, and contributing to blurring boundaries between expectations of the workplace and a worker's autonomy. We argue that this pattern would contribute to data colonialism [143] and we describe how the often-invisible failures of emotion recognition technology could influence high-stakes workplace decisions and exacerbate existing inequities.

6.1 Visibility Implications of Emotion Recognition in the Workplace

Introducing new technology into the workplace can alter the balance of power and the conditions of work by changing what is visible and what is hidden [6, 134, 140]. As our analysis shows, proposed emotion recognition systems claim to surface the internal states of their targets (i.e., workers) and report on their emotional states under circumstances where workplace display rules would normally *not* apply (such as when supervisors and customers are not nearby). We argue that the resulting increased visibility of emotional states as described in the analyzed patent applications could create additional labor for workers, compromise worker privacy, and contribute to a larger pattern of blurring boundaries between expectations of the workplace and a worker's autonomy. Emotion recognition technology's *failure*, on the other hand, is not especially visible to human decision makers. These shifts in visibility could have implications for design and use of emotion recognition, which we discuss below.

6.1.1 Additional Emotional Labor

Our findings reveal an additional implication of automated emotion recognition technology in the workplace beyond what prior work suggests (i.e., contested validity and potential for reifying societal biases [36]). In a workplace with emotional labor expectations where emotion recognition technology is introduced, workers would need to manage their emotional expressions in a way that can be read as appropriate both by the humans who evaluate their emotions and by an emotion recognition system. This resonates with recent research suggesting that social media users also engage in emotional labor as a result of being monitored by emotion recognition technology on social media [7].

Any difference between the expectations of the human and the algorithmic evaluators constitutes additional labor for all workers subject to automated emotion recognition. We argue that this impact is disproportionately high for workers who have higher emotional labor expectations (e.g., service workers and women) [138] and for workers for whom emotional management or expression is more challenging (for example, autistic workers [76, 98]). This compounds the fact that marginalized workers (e.g., along dimensions of race, ethnicity, gender, sexuality, disability) are historically likely to have more emotional labor expectations at work, be more closely surveilled, and have less privacy, than cisgender white men and other higher status workers [77].

Further, the higher the error rate the emotion recognition system has, the more difference is possible between the emotional behavior they need to exhibit to satisfy their human and algorithmic evaluators. Therefore, groups for which the system is less accurate get more additional emotional labor when emotion recognition technology is implemented than groups with higher accuracy rates. For example, machine learning systems that use facial images (25 patent applications in our corpus), text (40), and speech (37) have been shown to perform worse on women and people with dark skin [17, 21, 41, 97]. Facial expression detection software more frequently misclassifies subjects with darker skin as being angry, an emotional expression that is frequently specifically penalized by systems in this corpus [116].

Similarly, systems that include physical activity data (21), like motion, gestures, and posture, in their input data could be biased against people with physical disabilities.

The potential harm for marginalized groups is starkly revealed by imagining the implementation of patent applications that claim to determine criminal intent and can alert police. However, even in a future where the actions taken by emotion AI seems to have low stakes, like playing calming music, the implementation of that technology may disproportionately negatively impact marginalized individuals' conditions of work. For example, an application discussed above (0120219) uses facial images to assess workers' emotions on a factory floor and play music or video content to "calm or settle" a person when they are determined to be angry. If that technology consistently assesses a Black worker to be angry, it will repeatedly interrupt them, potentially hurting the quality of their work. Such interruptions can increase stress and anxiety, and requiring additional labor to manage the alerts and cope with the interruptions [49, 79, 152]. If other workers notice that the Black employee gets more alerts than the others, it may have social consequences, including reinforcing the very harmful bias that shaped the algorithms to begin with. Eventually, we speculate that the pattern of interruptions and social consequences could cause the worker to be angered by the "calming" intervention.

Making previously difficult-to-measure emotional labor more visible could create a future with more empowered workers by supporting their demanding compensation for that labor. Tracking and documenting work has been a strategy that low-wage and precarious workers used to address wage theft [44]. However, simply adding surveillance technology is unlikely to increase employee pay if employers do not acknowledge the additional work it imposes. It is possible that employers view emotional labor as part of the job to begin with and see emotion recognition technology as a perfect measurement tool (as they sometimes claim to be) – not an additional expectation or an accounting of something they were not previously aware of. Without a clear, organized argument about the additional emotional labor and other costs emotional surveillance imposes on workers, this emerging technology will likely fit a larger pattern of job scope creep resulting from increased data collection [73, 111] without compensation. Further, some of these technologies require targets to annotate data about their own emotion, potentially requiring them to relive emotional moments in their work day and adding additional work to keep up with the accounting of their labor.

To reduce the risks of an erroneous emotion estimate and avoid punishment, workers under automated emotion surveillance may be required to manage their emotional display so that it reads appropriately to customers, their supervisor, and an automated system that they may not understand and may be biased against them because of their skin color, cultural background, or disability. We advocate that guidelines for evaluating an emotion recognition system should include balancing claimed benefits against the additional emotional labor required from workers.

6.1.2 Additional Privacy Threats and Going Beyond "Face" Data

Our findings also enable speculation on potential futures of surveillance and privacy. Emotion recognition patent applications in our corpus frequently promise to detect and predict the emotions of a target person or group of people in circumstances better than another person can, or in circumstances when another person is not able to. People generally operate with the expectation that others can see only their external emotional expressions, not their *true internal* state. Typically, research suggests that workers use response modulations like neutralizing or masking emotions [87] for politeness, to serve organizational goals, to manage privacy, to protect themselves, or to deceive. Similarly, people living with mental illnesses rely heavily on response modulations to manage their conditions and comply with social norms while at work [87].

Many of the analyzed patent applications claim to pierce workers' external emotional expression and allow supervisors to see workers' *hidden*, *internal* states (a pattern most directly exemplified by patent 10410655, as described in section 5.2). If these technologies do work as claimed, employers who implement them would have a view into the *truly private*: the internal mental state of their workers. They would also be also just as likely to detect data resulting from work-related emotion as non-work-related ones. They may collapse polite fictions that workplace relationships depend on, like "I am not annoyed at my supervisor," or "this job is my highest priority." Many common feelings must be hidden to conform to norms of professionalism, like personal annoyance or sexual attraction. Consider this scenario: if a person can conceal these undesirable feelings at work well enough to prevent anyone from noticing those feelings and they go on to do their job, organizational goals are achieved and politeness is preserved. However, if the supervisor uses an emotion recognition technology like the ones described in our corpus to uncover that an employee is annoyed at their boss, for example, using the legitimation of being able to respond to employee feelings, the annoyance becomes visible to the boss. Perhaps the annoyance is based on a personal feature of the worker or manager, or it is out of the manager's ability to address. The annoyance is no longer secret, nor can it be mitigated, but it interferes in the relationship between the worker and manager once it is surfaced.

Even outside of a workplace context, many find emotion recognition technologies invasive and problematic; social media users, for example, associate the use of automated emotion recognition technology with a loss of autonomy and control [7] and students are concerned about the accuracy, validity, privacy, and security of affect data [153]. With their livelihood on the line, we would expect targets of emotion recognition to be more sensitive to accuracy, privacy, and autonomy threats – understanding workers' perspectives is a high priority for future work.

Because of the stigma against certain health conditions and the risk of disparate treatment at work, many countries, including the U.S., have special protections for information about workers' or candidates' medical conditions [48]. However, the outputs of emotion recognition systems may reveal symptoms of protected medical conditions and emotions and moods related to personal events that a worker could keep private from their co-workers, but not from always-on and sometimes surreptitious biometric sensors. This would constitute employers skirting around rules in place to protect workers' medical privacy and exposing them to potential stigma at work. In this future, emotion recognition not only encroaches on workers' privacy, it also reinforces existing power imbalances between workers and employers by virtue of employers having access to important information the workers may not have wished to disclose.

The theory of Contextual Integrity [105] asserts that privacy includes protection for information that is tailored to the norms of specific contexts. The patent applications in our corpus frequently describe technologies that do not respond to and are apparently designed without regard to context (for example, one technology being applied without alteration between parents and children, teachers and students, and supervisors and workers). They also frequently claim long lists of data that their system could potentially use, presumably to avoid limiting the scope of their patent or perhaps that is how these technologies are truly imagined. A patent evaluation and granting system, like the USPTO [56], that protects claims and evaluates for novelty, usefulness, and non-obviousness encourages broad claims and data collection without consideration for contextual integrity. Like other data-driven workplace technologies, the data collection and use of emotion recognition would require more transparency from the worker to the employer while further obscuring employers' arcane decision-making process by placing a black box in the center [5].

Many patent applications in our set describe using other data in addition to the *emotion* data to improve emotion estimates or target system actions, like demographic information, communication metadata, geographic location, schedules, environmental information, customer data, or business characteristics. Although broadening the scope of used input data could offer some useful context to avoid unwanted action (e.g., suggesting an action that would be inappropriate or impossible in the target's location), it could also increase the threat to workers' privacy. In addition to simply collecting and retaining more data, merging datasets can produce insights (and risks) that targets may not be able to predict [97]. The mythos that more data is always better, pervasive in the big data rhetoric, increases the incentive for broad ranging data collection without concern for targets' privacy expectations and attention to the social contexts within which technologies are to be deployed [34, 143]. If employed, we see this approach as likely to contribute to what scholars have called data colonialism: a normalization of capitalist exploitation based on diverse data types that celebrates "the availability of new data sources and the potential for new insights and perspectives they may enable" and "colonizes and commodifies everyday life in ways previously impossible" [143].

Finally, while *facial* emotion recognition was a common type of emotion recognition present in our corpus, the analyzed patent applications used a wide variety of data types and sources to infer intimate details about their targets. We appreciate and echo recent advocacy efforts to resist the harms of facial recognition technology more broadly [38], arguing that it violates rights to privacy and due process, can exacerbate existing biases in policing, and shields error- and bias-prone estimates a false sense of objectivity. However, our analysis also renders us concerned: we identified broad input data types and language in patent applications, which perhaps are intended to avoid tying patented technology to a particular data type or to control the future of the technology. Nonetheless, this broad scope is a political choice with implications; for example, if implemented, it may support employers, regulators, technologists, or other actors to move away from the collection of facial data and simply shift to other sensitive data types such as those we uncovered in our analysis without truly addressing their fundamental harmful implications. While attention to facial recognition's harms is absolutely crucial, we caution against attending to automated *emotion recognition* only insofar as it relies on and overlaps with *facial recognition,* noting that emotion recognition's input data goes beyond "the face." We hope this work contributes to shaping and facilitating public and regulatory critical attention regarding these other data types that have garnered less attention and are gaining more recent attention from actors such as the US OSTP as noted in the Introduction section.

Overall, we argue that emotion recognition technologies, if implemented as described in the applications we analyzed, could pose substantial privacy threats in the workplace and contribute to harmful trends in data colonialism [143].

6.1.3 Boundary Blurring

Emotion recognition technology as described in our corpus could contribute to a pattern of workplace technology blurring boundaries between a target's work, governed by their employers' policies and norms, and their own life, over which they have more autonomy [64, 99]. A boundary between work and life might exist at the door of a work building, until a laptop, cloud software, a VPN, or remote desktop software allow workers to work outside the office. A time boundary might exist between work and life, for example at 9:00 am and 5:00 pm, or that line can be blurred by a mobile phone or another device that allows an employee to be contacted outside of work. Tracking software in the workplace has consistently contributed to this kind of boundary blurring in the past [99].

Similar to geographic and time boundaries, a boundary may exist between an employee's work and life between the worker's external emotional expression and internal, experienced emotions. If emotion recognition software blurs this line, workers may lose autonomy over their internal state while working. Workers seem to experience less dissonance and stress when their internal states match their external expression [61]. However, the worker may feel additional pressure to allow work to take over areas, times, and thoughts that they previously had control over.

We speculate that power would likely complicate the boundary blurring shaped by emotion recognition in the workplace. Targets of the technology rely on the income from their work to provide for food and shelter and are therefore not fully free to opt out, alter, or undermine surveillance. The importance of power relationships and the social context of employment was frequently absent from the analyzed patent applications, which often described a range of use cases for the same technology across diverse contexts without acknowledgement of any potential differences.

Some of the analyzed applications include some features that could give more agency to emotion recognition technology's targets. For example, they may give the target the option to edit the emotions that were inferred, alter the system's default settings, or intervene in outputs on a case-by-case basis. However, if the target is an employee at work, policies or supervisors' instructions may limit the target's ability to exercise this agency, and instead those controls are effectively in the supervisor's hands. It is also possible that workers would feel the need to prove that they have nothing to hide (a common response in privacy and technology studies [133]), and therefore would not practice this agency to avoid the implication that they do have something to hide – creating additional labor and rendering control features useless for targets. While these are speculations on possible implications [53] and future work could directly engage with workers and employers, they do allow us to consider potential ethical implications of emotion recognition technologies in the workplace.

In summary, we argue that emotion recognition technology in the workplace could reduce worker autonomy by imposing workplace expectations on workers' internal states. This would follow and extend a larger pattern of employers using technology to blur boundaries and, in so doing, extend the reach of their expectations into an area where a worker might have otherwise had full autonomy: in this case, the workers' own internal states. Further, the rare guardrails that emotion recognition patent applications describe to protect worker autonomy can be compromised by power relationships (between the employer and worker) inherent in labor structures.

6.2 The (in)visibility of Emotion Recognition Technology's Failure and Implications for Technologists

We argue that by using opaque algorithms and failing to include transparency features or measures of uncertainty, however, the emotion recognition technologies described in this corpus would be likely to fail *invisibly*. Prior work illustrates some of the potential quality flaws (e.g., higher error rates in unconstrained conditions, biases in outcomes, and potential bias in error rates) leading to failure in the technology underlying automated emotion recognition systems. CSCW has long considered the failure of workplace technology [67], which very often fails in a visible way. For example, workers may not use the technology (e.g., [58, 147]), it may not be maintained (e.g., [25]), or may not be trusted [94]. In each of these cases, a human decision maker is seeing the failure and taking it into consideration.

First, an emotion recognition system may have a crucial flaw embedded in its design assumptions. Directly or implicitly, they assume connections between their input data and the estimates or predictions they make, raising validity questions: can the input data (e.g., heart rate or facial expression) estimate what they claim it can (e.g., emotion or mood)? Are the data fully representative of what the application claims they measure? Does knowing emotion support the actions that the system takes or prompts? However, if the applications in this corpus are any indication, these concerns do not seem to be accounted for in developing emotion recognition technologies or making arguments for patent requirements of novelty, non-obviousness, and usefulness. To support public interests being considered at a key moment in determining which technological futures we move into, we echo suggestions to lower expertise barriers and incorporate public participation in patent proceedings [159]. We suggest that the USPTO consider encouraging applicants to contend with questions regarding invisible failures in their applications by interpreting validity as part of the "usefulness" requirement, and that designers carefully consider raising the visibility of uncertainty estimates and potential failure points.

Second, emotion recognition technologies currently have nonzero, and often high, error rates [2, 96]. Several patent applications claimed that their estimations were objective, unbiased, accurate, or could be used to avoid human bias in emotion recognition. Research concerned with ethics in emotion recognition has emphasized the importance of not treating predictions as ground truth and clearly communicating the limitations of system's use [71]. Just six applications in this corpus included a measure of confidence in their reports to human decision makers or reported their results as a probability. In a future where error-prone emotion recognition technology without intelligible and actionable confidence reporting, the supervisors, mental health professionals, or authorities who get this data without context may make decisions based on shaky or false estimations with inappropriate confidence. This could result in workers being penalized, fired, or arrested based on emotions that they would have successfully hidden from a human decision-maker or did not even have in the first place.

Third, demographically uneven failure may cause more harm than a similar level of overall failure. For example, it is well-established that training data sets do not include sufficient data on Black Americans' faces, speech, or text [22, 40, 45, 65, 81]. If a manager sees that basing their decisions on their algorithm's output is working well *most of the time*, they may develop what would be an inaccurate or rather harmful trust in it. Consider this scenario: if the emotion recognition system *invisibly* fails more frequently for Black workers, but the manager is basing their trust of the system on the higher quality level of outputs for the larger group of workers, the manager may believe that their Black employee(s) are more often inappropriately angry, for example, causing the employees to be inappropriately punished and perpetuating existing racist stereotypes (e.g., [101]) and injustices. If implemented emotion recognition technologies have higher error rates for groups that are already underrepresented and underpaid, it could exacerbate existing harms in the workplace.

We argue that designers must consider the possibility and visibility of failure when designing emotion recognition technology. Making the uncertainty of emotion inferences highly visible could help decision makers, employers, and workers contextualize their response to the inferences. However, not all failure is visible through uncertainty reporting: validity threats are difficult to make visible. Additionally, a confidence estimate that may sound high to a human decision maker, like 97%, would still generate 3 failures for every one hundred estimates. If implemented in a large company and used in real-time, that could mean thousands of confident, false estimates every day. If a system can have thousands of daily confident, false estimates, basing decisions like hiring, firing, referring employees to security or law enforcement, reward, or punishment on single estimates would constitute injustice (and potential legal liability). If a system has higher error rates for some groups or individuals, even basing these decisions on a pattern of estimates is unreliable and harmful. Designers and users must both ensure that error rates can be and are taken into consideration when designing systems, deciding to employ them in the workplace, or acting on system outputs.

Transparency around training and input data could also raise the visibility of failure. Communicating clearly to the potential targets, regulators, and consumer watchdogs about the scope of training and input data collection is necessary for those groups to make informed decisions. Clarity around training processes is also a necessary step to achieve contestability [102].

Communicating to employers and other emotion recognition technology users how to collect and label training data for initial training, online learning, or personalization could help prevent some of the accidental inclusion of "garbage in" to the models. In any domain, it is not a straightforward or easy task to collect diverse training data that is close enough to the real-world use case, that oversamples rare cases where needed, and is accurately labeled. Training data collection can also ethical issues around as consent, positionality, power, contextuality, inclusivity, transparency, and privacy [71, 80, 98, 125] Emotion recognition training data is particularly difficult because: 1), humans, who may label such data, struggle to recognize emotions accurately [16]. Indeed, humans may struggle to recognize their own emotions accurately [7]; 2), appropriate training data diversity may be difficult to collect and impossible to induce; how do you make sure that you have enough examples of a range of emotions, including those that are not common at work? Algorithms and humans have exhibited bias in interpreting the emotional expressions of racial minority groups [22, 78, 116]: can you be sure you have enough examples across both emotion categories *and* demographic groups to ensure that your model does not render results or error rates that are biased against any group? How confident are you in your annotators' accuracy with data about transgender and non-binary people [126]? Can you collect such data, oversampling rare cases when necessary, without compounding existing disparities in surveillance at work? While we pose these questions for designers to consider, we emphasize that these questions are *only* relevant after it is established that emotion recognition *is* appropriate, valid, and respectful to targets – an assumption that recent work and debate has critiqued [7, 16, 36, 120, 135]. That is, before embarking on efforts to train data appropriately or diversifying datasets (which can lead to more surveillance and targeting of minoritized groups [41, 80], designers should ask if they *should* design emotion recognition technologies to begin with, and if diversifying training datasets with inclusion of more minoritized groups is something those groups desire and would benefit from on their *own* terms.

It is likely that patent applications are not the right place for this communication to any of these stakeholders. Nonetheless, our analysis highlighted the potentially harmful implications of users, targets, and others relying *uncritically* on the outputs of automated emotion recognition technology in the workplace and a concerning pattern of designers failing to include features to identify, report, or mitigate the possibility, potential consequences, and visibility of failure. As such, our findings emphasize the need for designers to reason seriously about the social impacts of their system [62], as failures of emotion recognition in the workplace can be invisible (difficult to recognize), high stakes, and can exacerbate inequalities and injustices that workers and job candidates face. We echo encouragements for designers of emotion recognition to carefully consider the risks posed by commercial applications of automated emotion recognition [71], the groups (including data subjects) it may impact [7, 107, 137] , and the ethical implications its framing, data collection, methods, and evaluation [98] *before* building, and to consider that there are situations in which these technologies should not be built at all [14, 20] if we consider their potentially harmful implications.

7 CONCLUSION

Patent applications give researchers, regulators, users, and citizens a view into an imagined future and the advertised claims about technology. We use patent applications to engage in ethical speculation about possible futures of and with emotion recognition technology. We contribute a landscape analysis of emotion recognition technology in the workplace by analyzing available patent applications (N=86). We identify these technologies' input data (along with associated validity concerns and training data and processes), promised outputs, and suggested actions based on outputs. We find that these technologies use a wide range of data, claim to reveal not only targets' emotional expressions, but also their internal states, and prompt a wide range of actions, many of which impact workers' employment and livelihoods. We find that the emotion recognition technologies being imagined by patent applicants often violate ethical guidelines suggested by researchers, especially their recommendations not to treat predictions as ground truth; not to claim to measure internal states; and to give data and decision subjects opportunities for feedback, customization, and recourse.

Within an ethical speculation framework, we argue that if implemented, these technologies increase the visibility of emotional states, have the potential to create additional labor for workers, can compromise worker privacy, and could contribute to a larger pattern of blurring boundaries between expectations of the workplace and a worker's autonomy, and more broadly to the data colonialism regime. The applications also reveal the potential for invisible failure in automated emotion recognition technology designs, which could inappropriately influence high-stakes workplace decisions and exacerbate inequities.

We recommend that companies exercise caution when designing emotion recognition technology and/or deploying it in the workplace. If they elect to implement emotion recognition in the workplace, we argue that they must consider validity, accuracy, and bias questions in data collection; the social implications of making workers' internal states more visible; how to manage the considerable privacy and autonomy threats posed by emotion recognition in the workplace; how to increase the visibility of potential and invisible technology failures, and what to do when these failures occur.

We recommend that emotion recognition technologies' designers carefully consider the potential for their technologies to contribute to emotional labor expectations for workers, disparate harmful impact, the implications of surfacing the supposed internal states of low-power workers, and the invisibility of failure. We emphasize the choice that designers have to *not* build automated emotion recognition technologies for use in high-stakes contexts like the workplace. If they chose to design such systems, we recommend that they give clear training to the humans who will use and be affected by them, that they call attention to uncertainty and failure potential, and that they give workers access to view, contextualize, or even edit data about themselves. We encourage designers to contend with the impacts of their training, data collection outputs, and encouraged actions given how difficult it is to do emotion recognition robustly (automatic or otherwise).

As workplace surveillance increases its reach beyond behavior and into the internal states of workers, we recommend that regulators provide protection for workers and job candidates and consider adding additional scrutiny to patent applications and subsequent technologies focused on their societal implications. We hope that this work helps shape conversations and decisions such as Federal calls posed by the U.S. OSTP, noted in our Introduction, seeking to better understand biometric technologies including emotion recognition and AI [106].

ACKNOWLEDGMENTS

The authors would like to thank Shobita Parthasarathy for her comments on this work and Lenna Lalwani for her help scoping our patent queries. This work was supported by the National Science Foundation (award number: 2020872).

REFERENCES

[1] Abbas, A. et al. 2014. A literature review on the state-of-the-art in patent analysis. *World Patent Information*. 37, (Jun. 2014), 3–13. DOI:https://doi.org/10.1016/j.wpi.2013.12.006.

[2] Ahmed, F. et al. 2020. Emotion Recognition From Body Movement. *IEEE Access*. 8, (2020), 11761–11781. DOI:https://doi.org/10.1109/ACCESS.2019.2963113.

[3] Aiello, J.R. and Kolb, K.J. Electronic Performance Monitoring and Social Context: Impact on Productivity and Stress. 15.

[4] Ajunwa, I. et al. 2017. Limitless Worker Surveillance. (2017). DOI:https://doi.org/10.15779/Z38BR8MF94.

[5] Ajunwa, I. 2020. The "black box" at work. *Big Data & Society*. 7, 2 (Jul. 2020), 2053951720966181. DOI:https://doi.org/10.1177/2053951720938093.

[6] AM CSCW Challenging Perspectives on Work and Technology: 1994. *https://www.semanticscholar.org/paper/AM-CSCW-Challenging-Perspectives-on-Work-and-Bannon/656e0ff165c4eb411435eedfda9d098ff8544807*. Accessed: 2021-07-10.

[7] Andalibi, N. and Buss, J. 2020. The Human in Emotion Recognition on Social Media: Attitudes, Outcomes, Risks. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu HI USA, Apr. 2020), 1–16.
[8] Andalibi, N. and Forte, A. 2018. Announcing Pregnancy Loss on Facebook: A Decision-Making Framework for Stigmatized Disclosures on Identified Social Network Sites. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery. 1–14.
[9] Andrade, E.B. and Ariely, D. 2009. The enduring impact of transient emotions on decision making. *Organizational Behavior and Human Decision Processes*. 109, 1 (May 2009), 1–8. DOI:https://doi.org/10.1016/j.obhdp.2009.02.003.
[10] Angie, A.D. et al. 2011. The influence of discrete emotions on judgement and decision-making: A meta-analytic review. *Cognition and Emotion*. 25, 8 (Dec. 2011), 1393–1422. DOI:https://doi.org/10.1080/02699931.2010.550751.
[11] Auerbach, C. and Silverstein, L.B. 2003. *Qualitative Data: An Introduction to Coding and Analysis*. NYU Press.
[12] Ball, K. 2010. Workplace surveillance: an overview. *Labor History*. 51, 1 (Feb. 2010), 87–106. DOI:https://doi.org/10.1080/00236561003654776.
[13] Bard, J. 2020. *Developing a Legal Framework for Regulating Emotion AI*. Technical Report #3680909. Social Science Research Network.
[14] Barocas, S. et al. 2020. When not to design, build, or deploy. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (New York, NY, USA, Jan. 2020), 695.
[15] Barrett, L.F. et al. 2011. Context in Emotion Perception. *Current Directions in Psychological Science*. 20, 5 (Oct. 2011), 286–290. DOI:https://doi.org/10.1177/0963721411422522.
[16] Barrett, L.F. et al. 2019. Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements. *Psychological Science in the Public Interest*. 20, 1 (Jul. 2019), 1–68. DOI:https://doi.org/10.1177/1529100619832930.
[17] Barrett, L.F. et al. 2001. Knowing what you're feeling and knowing what to do about it: Mapping the relation between emotion differentiation and emotion regulation. *Cognition and Emotion*. 15, 6 (Nov. 2001), 713–724. DOI:https://doi.org/10.1080/02699930143000239.
[18] Barrett, L.F. 2017. The theory of constructed emotion: an active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience*. 12, 1 (Jan. 2017), 1–23. DOI:https://doi.org/10.1093/scan/nsw154.
[19] Barsade, S.G. and Gibson, D.E. 2007. Why Does Affect Matter in Organizations? *Academy of Management Perspectives*. 21, 1 (Feb. 2007), 36–59. DOI:https://doi.org/10.5465/amp.2007.24286163.
[20] Baumer, E.P.S. and Silberman, M.S. 2011. When the implication is not to design (technology). *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver BC Canada, May 2011), 2271–2274.
[21] Bazarova, N.N. et al. 2015. Social Sharing of Emotions on Facebook: Channel Differences, Satisfaction, and Replies. *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing* (New York, NY, USA, Feb. 2015), 154–164.
[22] Blodgett, S.L. and O'Connor, B. 2017. Racial Disparity in Natural Language Processing: A Case Study of Social Media African-American English. *arXiv:1707.00061 [cs]*. (Jun. 2017).
[23] Boldrin, M. and Levine, D.K. 2013. The Case Against Patents. *Journal of Economic Perspectives*. 27, 1 (Feb. 2013), 3–22. DOI:https://doi.org/10.1257/jep.27.1.3.
[24] Bolukbasi, T. et al. 2016. Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings. *Advances in Neural Information Processing Systems 29*. D.D. Lee et al., eds. Curran Associates, Inc. 4349–4357.
[25] Bowers, J. 1994. The work to make a network work: studying CSCW in action. *Proceedings of the 1994 ACM conference on Computer supported cooperative work* (New York, NY, USA, Oct. 1994), 287–298.
[26] Brotheridge, C.M. and Grandey, A.A. 2002. Emotional Labor and Burnout: Comparing Two Perspectives of "People Work." *Journal of Vocational Behavior*. 60, 1 (2002), 17–39.
[27] Bullington, J. 2005. "Affective" computing and emotion recognition systems: the future of biometric surveillance? (Sep. 2005), 95–99. DOI:https://doi.org/10.1145/1107622.1107644.
[28] Buolamwini, J. and Gebru, T. 2018. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. (2018), 15.
[29] Calero Valdez, A. and Ziefle, M. 2018. Human Factors in the Age of Algorithms. Understanding the Human-in-the-loop Using Agent-Based Modeling. *Social Computing and Social Media. Technologies and Analytics* (Cham, 2018), 357–371.
[30] Caliskan, A. et al. 2017. Semantics derived automatically from language corpora contain human-like biases. *Science*. 356, 6334 (Apr. 2017), 183–186. DOI:https://doi.org/10.1126/science.aal4230.
[31] Chancellor, S. et al. 2019. A Taxonomy of Ethical Tensions in Inferring Mental Health States from Social Media. *Proceedings of the Conference on Fairness, Accountability, and Transparency* (Atlanta GA USA, Jan. 2019), 79–88.

[32] Chancellor, S. et al. 2019. The Relationships between Data, Power, and Justice in CSCW Research. *Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing* (New York, NY, USA, Nov. 2019), 102–105.

[33] Corvite, S. et al. 2022. Data Subjects' Perspectives on Emotion Artificial Intelligence Use in the Workplace: A Relational Ethics Lens. (2022).

[34] Couldry, N. and Mejias, U.A. 2019. *The Costs of Connection: How Data Is Colonizing Human Life and Appropriating It for Capitalism*. Stanford University Press.

[35] Cranor, L.F. A Framework for Reasoning About the Human in the Loop. 15.

[36] Crawford, K. et al. 2019. *AI Now 2019 Report*.

[37] Crawford, K. 2021. *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press.

[38] Crawford, K. 2019. Halt the use of facial-recognition technology until it is regulated. *Nature*. 572, 7771 (Aug. 2019), 565–565. DOI:https://doi.org/10.1038/d41586-019-02514-7.

[39] Daim, T.U. et al. 2006. Forecasting emerging technologies: Use of bibliometrics and patent analysis. *Technological Forecasting and Social Change*. 73, 8 (Oct. 2006), 981–1012. DOI:https://doi.org/10.1016/j.techfore.2006.04.004.

[40] Diao, J.A. and Adamson, A.S. 2021. Representation and misdiagnosis of dark skin in a large-scale visual diagnostic challenge. *Journal of the American Academy of Dermatology*. (Apr. 2021). DOI:https://doi.org/10.1016/j.jaad.2021.03.088.

[41] D'Ignazio, C. and Klein, L. 2020. 1. The Power Chapter. *Data Feminism*.

[42] D'Mello, S. and Calvo, R.A. 2013. Beyond the basic emotions: what should affective computing compute? *CHI '13 Extended Abstracts on Human Factors in Computing Systems* (New York, NY, USA, Apr. 2013), 2287–2294.

[43] Dobrosovestnova, A. et al. 2021. Service robots for affective labor: a sociology of labor perspective. *AI & SOCIETY*. (Apr. 2021). DOI:https://doi.org/10.1007/s00146-021-01208-x.

[44] Dombrowski, L. et al. 2017. Low-Wage Precarious Workers' Sociotechnical Practices Working Towards Addressing Wage Theft. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery. 4585–4598.

[45] Dorn, R. 2019. Dialect-Specific Models for Automatic Speech Recognition of African American Vernacular English. *Proceedings of the Student Research Workshop Associated with RANLP 2019* (Varna, Bulgaria, Sep. 2019), 16–20.

[46] Ekman, P. 1992. Are there basic emotions? *Psychological Review*. 99, 3 (1992), 550–553. DOI:https://doi.org/10.1037/0033-295X.99.3.550.

[47] Elfenbein, H.A. et al. 2007. Reading your Counterpart: The Benefit of Emotion Recognition Accuracy for Effectiveness in Negotiation. *Journal of Nonverbal Behavior*. 31, 4 (Dec. 2007), 205–223. DOI:https://doi.org/10.1007/s10919-007-0033-7.

[48] Employers and Health Information in the Workplace: 2008. *https://www.hhs.gov/hipaa/for-individuals/employers-health-information-workplace/index.html*. Accessed: 2021-07-06.

[49] Eyrolle, H. and Cellier, J.-M. 2000. The effects of interruptions in work activity: Field and laboratory results. *Applied Ergonomics*. (2000).

[50] Features | CompanyMood: *https://www.company-mood.com/features*. Accessed: 2021-07-06.

[51] Fenton-O'Creevy, M. et al. 2011. Thinking, feeling and deciding: The influence of emotions on the decision making and performance of traders. *Journal of Organizational Behavior*. 32, 8 (2011), 1044–1061. DOI:https://doi.org/10.1002/job.720.

[52] Fiesler, C. 2019. Ethical Considerations for Research Involving (Speculative) Public Data. *Proceedings of the ACM on Human-Computer Interaction*. 3, GROUP (Dec. 2019), 1–13. DOI:https://doi.org/10.1145/3370271.

[53] Fiesler, C. 2021. *Innovating Like an Optimist, Preparing Like a Pessimist: Ethical Speculation and the Legal Imagination*. Technical Report #ID 3779036. Social Science Research Network.

[54] Gartner: Workplace Analytics Needs Digital Ethics: 2015. *https://www.gartner.com/document/3164927?ref=authrightrec&refval=3887785*. Accessed: 2021-02-20.

[55] Gaudine, A. and Thorne, L. 2001. Emotion and Ethical Decision-Making in Organizations. *Journal of Business Ethics*. 31, 2 (May 2001), 175–187. DOI:https://doi.org/10.1023/A:1010711413444.

[56] General information concerning patents: *https://www.uspto.gov/patents/basics/general-information-patents*. Accessed: 2021-07-09.

[57] Gonçalves, A.R. et al. 2018. Effects of age on the identification of emotions in facial expressions: a meta-analysis. *PeerJ*. 6, (Jul. 2018), e5278. DOI:https://doi.org/10.7717/peerj.5278.

[58] Gorm, N. and Shklovski, I. 2016. Steps, Choices and Moral Accounting: Observations from a Step-Counting Campaign in the Workplace. *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing* (San Francisco California USA, Feb. 2016), 148–159.

[59] Gorrostieta, C. et al. 2019. Gender De-Biasing in Speech Emotion Recognition. *Interspeech 2019* (Sep. 2019), 2823–2827.

[60] Grandey, A.A. et al. 2015. Emotional labor threatens decent work: A proposal to eradicate emotional display rules. *Journal of Organizational Behavior*. 36, 6 (2015), 770–785. DOI:https://doi.org/10.1002/job.2020.
[61] Grandey, A.A. 2003. When "The Show Must Go On": Surface Acting and Deep Acting as Determinants of Emotional Exhaustion and Peer-Rated Service Delivery. *Academy of Management Journal*. 46, 1 (Feb. 2003), 86–96. DOI:https://doi.org/10.5465/30040678.
[62] Green, B. 2019. "Good" isn't Good Enough.
[63] Greene, G. 2020. *The Ethics of AI and Emotional Intelligence: Data sources, applications, and questions for evaluating ethics risk*. Partnership on AI.
[64] Gregg, M. 2013. *Work's Intimacy*. John Wiley & Sons.
[65] Groh, M. et al. 2021. Evaluating Deep Neural Networks Trained on Clinical Images in Dermatology With the Fitzpatrick 17k Dataset. (2021), 1820–1828.
[66] Grudin, J. 1994. Computer-supported cooperative work: history and focus. *Computer*. 27, 5 (May 1994), 19–26. DOI:https://doi.org/10.1109/2.291294.
[67] Grudin, J. 1988. Why CSCW applications fail: problems in the design and evaluationof organizational interfaces. *Proceedings of the 1988 ACM conference on Computer-supported cooperative work* (New York, NY, USA, Jan. 1988), 85–93.
[68] Guy, M.E. and Newman, M.A. 2004. Women's Jobs, Men's Jobs: Sex Segregation and Emotional Labor. *Public Administration Review*. 64, 3 (2004), 289–298. DOI:https://doi.org/10.1111/j.1540-6210.2004.00373.x.
[69] Harwell, D. A face-scanning algorithm increasingly decides whether you deserve the job. *Washington Post*.
[70] Harwell, D. Rights group files federal complaint against AI-hiring firm HireVue, citing 'unfair and deceptive' practices. *Washington Post*.
[71] Hernandez, J. et al. 2021. Guidelines for Assessing and Minimizing Risks of Emotion Recognition Applications. *2021 9th International Conference on Affective Computing and Intelligent Interaction (ACII)* (Sep. 2021), 1–8.
[72] Hochschild, A.R. 2012. *The managed heart: commercialization of human feeling*. University of California Press.
[73] Holten Møller, N.L. et al. 2017. Data Tracking in Search of Workflows. *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (New York, NY, USA, Feb. 2017), 2153–2165.
[74] How Emotion-Sensing Technology Can Reshape the Workplace: *https://sloanreview.mit.edu/article/how-emotion-sensing-technology-can-reshape-the-workplace/*. Accessed: 2021-07-06.
[75] Howard, A. et al. 2017. Addressing bias in machine learning algorithms: A pilot study on emotion recognition for intelligent systems. *2017 IEEE Workshop on Advanced Robotics and its Social Impacts (ARSO)* (Austin, TX, USA, Mar. 2017), 1–7.
[76] Hurley-Hanson, A.E. et al. 2020. The Stigma of Autism. *Autism in the Workplace: Creating Positive Employment and Career Outcomes for Generation A*. A.E. Hurley-Hanson et al., eds. Springer International Publishing. 21–45.
[77] Igo, S.E. 2018. *The Known Citizen*. Harvard University Press.
[78] Jack, R.E. et al. 2009. Cultural Confusions Show that Facial Expressions Are Not Universal. *Current Biology*. 19, 18 (Sep. 2009), 1543–1548. DOI:https://doi.org/10.1016/j.cub.2009.07.051.
[79] Jett, Q.R. 2003. Work Interrupted: A Closer Look at the Role of Interruptions in Organizational Life. *Academy of Management Review*. (2003), 14.
[80] Jo, E.S. and Gebru, T. 2020. Lessons from archives: strategies for collecting sociocultural data in machine learning. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (New York, NY, USA, Jan. 2020), 306–316.
[81] Kamulegeya, L.H. et al. 2019. Using artificial intelligence on dermatology conditions in Uganda: A case for diversity in training data sets for machine learning. *bioRxiv*. (Oct. 2019), 826057. DOI:https://doi.org/10.1101/826057.
[82] Kaur, H. et al. 2022. "I Didn't Know I Looked Angry": Characterizing Observed Emotion and Reported Affect at Work. *CHI Conference on Human Factors in Computing Systems* (New Orleans LA USA, Apr. 2022), 1–18.
[83] Khalil, S. 2014. Not everything that counts can be counted and not everything that can be counted counts. *The Psychiatric Bulletin*. 38, 2 (Apr. 2014), 86–86. DOI:https://doi.org/10.1192/pb.38.2.86b.
[84] Kim, Y.G. et al. 2008. Visualization of patent analysis for emerging technology. *Expert Systems with Applications*. 34, 3 (Apr. 2008), 1804–1812. DOI:https://doi.org/10.1016/j.eswa.2007.01.033.
[85] Konrath, S. et al. 2014. The Relationship Between Narcissistic Exploitativeness, Dispositional Empathy, and Emotion Recognition Abilities. *Journal of Nonverbal Behavior*. 38, 1 (Mar. 2014), 129–143. DOI:https://doi.org/10.1007/s10919-013-0164-y.
[86] Koops, B.-J. 2021. The concept of function creep. *Law, Innovation and Technology*. 13, 1 (Jan. 2021), 29–56. DOI:https://doi.org/10.1080/17579961.2021.1898299.

[87] Lawrence, S. et al. 2011. A Review of Emotion Regulation and Development of a Framework for Emotion Regulation in the Workplace. *Research in Occupational Stress and Well-being*. 197–263.

[88] Lee, Y.-R. et al. 2019. A Comparative Study of Burnout, Stress, and Resilience among Emotional Workers. *Psychiatry Investigation*. 16, 9 (Sep. 2019), 686–694. DOI:https://doi.org/10.30773/pi.2019.07.10.

[89] Leidner, R. 1999. Emotional Labor in Service Work. *The ANNALS of the American Academy of Political and Social Science*. 561, 1 (Jan. 1999), 81–95. DOI:https://doi.org/10.1177/000271629956100106.

[90] Lewig, K.A. and Dollard, M.F. 2003. Emotional dissonance, emotional exhaustion and job satisfaction in call centre workers. *European Journal of Work and Organizational Psychology*. 12, 4 (Dec. 2003), 366–392. DOI:https://doi.org/10.1080/13594320344000200.

[91] Lutchyn, Y. et al. 2015. MoodTracker: Monitoring collective emotions in the workplace. *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)* (Xi'an, China, Sep. 2015), 295–301.

[92] Mansfield, E. 1986. Patents and Innovation: An Empirical Study. *Management Science*. (Feb. 1986). DOI:https://doi.org/10.1287/mnsc.32.2.173.

[93] May, T. and Chen, A.C. 2021. "Smart" Office Cushions Track Workers by the Seat of Their Pants - The New York Times. *The New York Times*.

[94] Mcknight, D.H. et al. 2011. Trust in a specific technology: An investigation of its components and measures. *ACM Transactions on Management Information Systems*. 2, 2 (Jul. 2011), 12:1-12:25. DOI:https://doi.org/10.1145/1985347.1985353.

[95] Meditation app Headspace on track to double corporate clients, bring mindfulness to work: 2018. *https://www.cnbc.com/2018/09/02/companies-are-turning-to-headspace-to-help-their-workers-meditate.html*. Accessed: 2021-07-06.

[96] Mehta, D. et al. 2018. Facial Emotion Recognition: A Survey and Real-World User Experiences in Mixed Reality. *Sensors*. 18, 2 (Feb. 2018), 416. DOI:https://doi.org/10.3390/s18020416.

[97] Metcalf, J. et al. 2016. *Perspectives on Big Data, Ethics, and Society*. Council for Big Data, Ethics, and Society.

[98] Mohammad, S.M. 2022. Ethics Sheet for Automatic Emotion Recognition and Sentiment Analysis. *arXiv:2109.08256 [cs]*. (Mar. 2022).

[99] Møller, N.L.H. et al. 2017. A Constructive-Critical Approach to the Changing Workplace and its Technologies. (2017). DOI:https://doi.org/10.18420/ECSCW2017_P2.

[100] Moodbit | Employee Engagement: *https://mymoodbit.com/*. Accessed: 2021-07-06.

[101] Motro, D. et al. 2021. Race and reactions to women's expressions of anger at work: Examining the effects of the "angry Black woman" stereotype. *Journal of Applied Psychology*. (2021), No Pagination Specified-No Pagination Specified. DOI:https://doi.org/10.1037/apl0000884.

[102] Mulligan, D.K. et al. 2019. *Shaping Our Tools: Contestability as a Means to Promote Responsible Algorithmic Decision Making in the Professions*. Technical Report #ID 3311894. Social Science Research Network.

[103] Mustaqeem et al. 2020. Clustering-Based Speech Emotion Recognition by Incorporating Learned Features and Deep BiLSTM. *IEEE Access*. 8, (2020), 79861–79875. DOI:https://doi.org/10.1109/ACCESS.2020.2990405.

[104] Nepal, S. et al. 2022. A Survey of Passive Sensing in the Workplace. *arXiv:2201.03074 [cs]*. (Jan. 2022).

[105] Nissenbaum, H. 2004. Privacy as contextual integrity. *Wash. L. Rev.* 79, (2004), 119.

[106] Notice of Request for Information (RFI) on Public and Private Sector Uses of Biometric Technologies: 2021. *https://www.federalregister.gov/documents/2021/10/08/2021-21975/notice-of-request-for-information-rfi-on-public-and-private-sector-uses-of-biometric-technologies*. Accessed: 2021-12-26.

[107] Ong, D.C. 2021. An Ethical Framework for Guiding the Development of Affectively-Aware Artificial Intelligence. *2021 9th International Conference on Affective Computing and Intelligent Interaction (ACII)* (Sep. 2021), 1–8.

[108] Ortony, A. 2022. Are All "Basic Emotions" Emotions? A Problem for the (Basic) Emotions Construct. *Perspectives on Psychological Science*. 17, 1 (Jan. 2022), 41–61. DOI:https://doi.org/10.1177/1745691620985415.

[109] Parthasarathy, S. 2017. *Patent Politics: Life Forms, Markets, and the Public Interest in the United States and Europe*. University of Chicago Press.

[110] Picard, R.W. 2003. Affective computing: challenges. *International Journal of Human-Computer Studies*. 59, 1 (Jul. 2003), 55–64. DOI:https://doi.org/10.1016/S1071-5819(03)00052-1.

[111] Pine, K. 2016. The Work of Reuse: Birth Certificate Data and Healthcare Accountability Measurements. *iConference 2016 Proceedings* (Philadelphia, USA, Mar. 2016).

[112] Pugliesi, K. 1999. The Consequences of Emotional Labor: Effects on Work Stress, Job Satisfaction, and Well-Being. *Motivation and Emotion*. 23, 2 (Jun. 1999), 125–154. DOI:https://doi.org/10.1023/A:1021329112679.

[113] Quigley, K.S. et al. 2014. Inducing and measuring emotion and affect: Tips, tricks, and secrets. *Handbook of research methods in social and personality psychology, 2nd ed*. Cambridge University Press. 220–252.

[114] Raghavan, M. et al. 2020. Mitigating Bias in Algorithmic Hiring: Evaluating Claims and Practices. (2020), 13.

[115] Rahwan, I. 2018. Society-in-the-loop: programming the algorithmic social contract. *Ethics and Information Technology*. 20, 1 (Mar. 2018), 5–14. DOI:https://doi.org/10.1007/s10676-017-9430-8.
[116] Rhue, L. 2018. *Racial Influence on Automated Perceptions of Emotions*. Technical Report #ID 3281765. Social Science Research Network.
[117] Rimé, B. 2009. Emotion Elicits the Social Sharing of Emotion: Theory and Empirical Review. *Emotion Review*. 1, 1 (Jan. 2009), 60–85. DOI:https://doi.org/10.1177/1754073908097189.
[118] Roemmich, K. et al. 2023. Emotion AI at Work: Implications for Workplace Surveillance, Emotional Labor, and Emotional Privacy. *CHI '23: ACM Conference on Human Factors in Computing Systems* (2023).
[119] Roemmich, K. et al. 2023. Values in Emotion Artificial Intelligence Hiring Services: Technosolutions to Organizational Problems. (2023).
[120] Roemmich, K. and Andalibi, N. 2021. Data Subjects' Conceptualizations of and Attitudes toward Automatic Emotion Recognition-enabled Wellbeing Interventions on Social Media. (2021), 34.
[121] Russell, J.A. and Barrett, L.F. 1999. Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology*. 76, 5 (1999), 805–819. DOI:https://doi.org/10.1037/0022-3514.76.5.805.
[122] Sampat, B. 2018. *A survey of empirical evidence on patents and*. Technical Report #Working Paper 25383. National Bureau of Economic Research.
[123] Sandhu, N. et al. 2020. Human Emotions Detection Using Hybrid CNN Approach. *International Journal of Computer Science and Mobile Computing*. 9, 10 (Oct. 2020), 1–9. DOI:https://doi.org/10.47760/IJCSMC.2020.v09i10.001.
[124] Santamaria, K.Y. Facial Recognition Technology and Law Enforcement: Select Constitutional Considerations. 31.
[125] Scheuerman, M.K. et al. 2021. Do Datasets Have Politics? Disciplinary Values in Computer Vision Dataset Development. *Proceedings of the ACM on Human-Computer Interaction*. 5, CSCW2 (Oct. 2021), 317:1-317:37. DOI:https://doi.org/10.1145/3476058.
[126] Scheuerman, M.K. et al. 2019. How Computers See Gender: An Evaluation of Gender Classification in Commercial Facial Analysis Services. *Proceedings of the ACM on Human-Computer Interaction*. 3, CSCW (Nov. 2019), 1–33. DOI:https://doi.org/10.1145/3359246.
[127] Schumaker, R.P. et al. 2016. Predicting wins and spread in the Premier League using a sentiment analysis of twitter. *Decision Support Systems*. 88, (Aug. 2016), 76–84. DOI:https://doi.org/10.1016/j.dss.2016.05.010.
[128] Selinger, E. and Hartzog, W. 2019. Opinion | What Happens When Employers Can Read Your Facial Expressions? *The New York Times*.
[129] Seo, M.-G. et al. 2004. The Role of Affective Experience in Work Motivation. *Academy of Management Review*. 29, 3 (Jul. 2004), 423–439. DOI:https://doi.org/10.5465/amr.2004.13670972.
[130] Shaburov, V. and Monastyrshyn, Y. 2019. Emotion recognition for workforce analytics. US10496947B1. Dec. 3, 2019.
[131] Shih, P. 2018. Beyond Human-in-the-Loop: Empowering End-Users with Transparent Machine Learning. 37–54.
[132] So, J. et al. 2015. The psychology of appraisal: Specific emotions and decision-making. *Journal of Consumer Psychology*. 25, 3 (Jul. 2015), 359–371. DOI:https://doi.org/10.1016/j.jcps.2015.04.003.
[133] Solove, D.J. 2007. I've Got Nothing to Hide and Other Misunderstandings of Privacy. *San Diego Law Review*. 44, (2007), 745.
[134] Star, S.L. and Strauss, A. 1999. Layers of Silence, Arenas of Voice: The Ecology of Visible and Invisible Work. *Computer Supported Cooperative Work*. 8, 1–2 (Mar. 1999), 9–30. DOI:http://dx.doi.org.proxy.lib.umich.edu/10.1023/A:1008651105359.
[135] Stark, L. et al. 2020. "I Don't Want Someone to Watch Me While I'm Working": Gendered Views of Facial Recognition Technology in Workplace Surveillance. *Journal of the Association for Information Science and Technology*. 71, 9 (2020), 1074–1088. DOI:https://doi.org/10.1002/asi.24342.
[136] Stark, L. 2016. The emotional context of information privacy. *The Information Society*. 32, 1 (Jan. 2016), 14–27. DOI:https://doi.org/10.1080/01972243.2015.1107167.
[137] Stark, L. and Hoey, J. 2021. The Ethics of Emotion in Artificial Intelligence Systems. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (New York, NY, USA, Mar. 2021), 782–793.
[138] Steinberg, R. and Figart, D. 1999. Emotional Labor SinceThe Managed Heart. *Annals of The American Academy of Political and Social Science - ANN AMER ACAD POLIT SOC SCI*. 561, (Jan. 1999), 8–26. DOI:https://doi.org/10.1177/000271629956100101.
[139] Subhashini, R. and Niveditha, P.R. 2015. Analyzing and Detecting Employee's Emotion for Amelioration of Organizations. *Procedia Computer Science*. 48, (Jan. 2015), 530–536. DOI:https://doi.org/10.1016/j.procs.2015.04.131.
[140] Suchman, L. 1995. Making work visible. *Communications of the ACM*. 38, 9 (Sep. 1995), 56–64. DOI:https://doi.org/10.1145/223248.223263.

[141] Tag, B. et al. 2021. Making Sense of Emotion-Sensing: Workshop on Quantifying Human Emotions. *Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers*. Association for Computing Machinery. 226–229.
[142] Tatman, R. 2017. Gender and Dialect Bias in YouTube's Automatic Captions. *Proceedings of the First ACL Workshop on Ethics in Natural Language Processing* (Valencia, Spain, Apr. 2017), 53–59.
[143] Thatcher, J. et al. 2016. Data colonialism through accumulation by dispossession: New metaphors for daily data. *Environment and Planning D: Society and Space*. 34, 6 (Dec. 2016), 990–1006. DOI:https://doi.org/10.1177/0263775816633195.
[144] The rise of workplace spying: *https://theweek.com/articles/564263/rise-workplace-spying*. Accessed: 2021-07-06.
[145] Thompson, N.C. et al. 2020. The Computational Limits of Deep Learning. *arXiv:2007.05558 [cs, stat]*. (Jul. 2020).
[146] Tian, L. et al. 2022. *Applied Affective Computing*. Association for Computing Machinery.
[147] Treem, J.W. Technology Non-Use as Avoiding Accountability. 4.
[148] Verma, G.K. and Tiwary, U.S. 2017. Affect representation and recognition in 3D continuous valence–arousal–dominance space. *Multimedia Tools and Applications*. 76, 2 (Jan. 2017), 2159–2183. DOI:https://doi.org/10.1007/s11042-015-3119-y.
[149] Vlachostergiou, A. et al. 2014. Investigating Context Awareness of Affective Computing Systems: A Critical Approach. *Procedia Computer Science*. 39, (Jan. 2014), 91–98. DOI:https://doi.org/10.1016/j.procs.2014.11.014.
[150] Voice Catchers | Yale University Press: *https://yalebooks.yale.edu/book/9780300248036/voice-catchers*. Accessed: 2021-12-26.
[151] Volkova, S. et al. 2013. Exploring Demographic Language Variations to Improve Multilingual Sentiment Analysis in Social Media. *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing* (Seattle, Washington, USA, Oct. 2013), 1815–1827.
[152] Wajcman, J. and Rose, E. 2011. Constant Connectivity: Rethinking Interruptions at Work. *Organization Studies*. 32, 7 (Jul. 2011), 941–961. DOI:https://doi.org/10.1177/0170840611410829.
[153] Wang, Q. et al. Sensing Affect to Empower Students: Learner Perspectives on Affect-Sensitive Technology in Large Educational Contexts. 14.
[154] Watkins Allen, M. et al. 2007. Workplace Surveillance and Managing Privacy Boundaries. *Management Communication Quarterly*. 21, 2 (Nov. 2007), 172–200. DOI:https://doi.org/10.1177/0893318907306033.
[155] Yan, G. et al. 2014. A bilingual approach for conducting Chinese and English social media sentiment analysis. *Computer Networks*. 75, (Dec. 2014). DOI:https://doi.org/10.1016/j.comnet.2014.08.021.
[156] Zenonos, A. et al. 2016. HealthyOffice: Mood recognition at work using smartphones and wearable sensors. *2016 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)* (Mar. 2016), 1–6.
[157] 2021. *Emotional Wellness Will Rescue Your Organization and Distributed Workforce*. Maverick* Research.
[158] 2019. *Manual of Patent Examining Procedure*.
[159] 2022. Opinion | Save America's Patent System. *The New York Times*.
[160] The New Future of Work. *Microsoft Research*.

APPENDICIES

A.1 Patent Assignees

Assignee	*Number*	*Country*
IBM Corp	16	USA
Samsung Electronic Co Ltd	3	South Korea
Sony Corp	3	Japan
24 7 AI Inc	2	USA
Adobe	2	USA
Avaya Inc	2	USA

Hitachi	2	Japan
Intel Corp	2	USA
Panasonic IP Mgmt Co	2	Japan
SAP SE	2	Germany
Sensory Logic Inc	2	USA
Wipro Ltd	2	India
AABBYY Prod LLC	1	USA
AT & T IP I LP	1	USA
Automatic Data Processing Inc	1	USA
BBY Solutions (Best Buy)	1	USA
Cisco Tech Inc	1	USA
Conduent Business Services	1	USA
Emerging Automotive LLC	1	USA
FMR LLC	1	USA
Fronteo Inc	1	Japan
Fuji Xerox Co Ltd	1	Japan
Fujitsu Ltd	1	Japan
Fuvi Cognitive Network Corp	1	USA
Genesys Telecom Lab Inc	1	USA
Hon Hai PREC IND CO LTD	1	Taiwan
HU MAN REN GONG ZHI NENG KE JI SHANGHAI LTD	1	China
IMATEC INC	1	The Netherlands
Jabfab Inc	1	USA
Japan Mathematical Inst Inc	1	Japan
Japan Science & Tech Agency	1	Japan
Johnson & Johnson Health & Wellness Solutions Inc	1	USA
JP Morgan Chase Bank	1	USA
Kanjoya Inc	1	USA
Lear Corp	1	USA
Newvoicemedia Ltd	1	England
Next Jump Inc	1	USA
Norton LifeLock Inc	1	USA

Nuance Comm Inc	1	USA
Nuralogix	1	Canada
NVISO SARL	1	Switzerland
Omron Corp	1	Japan
Qualtrics LLC	1	USA
Sestek Ses Ve Iletisim Bilgisayar Teknolojileri Sanayii Ve Ticaret As	1	Turkey
Slomkowski, Robin S	1	USA
Snap Inc	1	USA
Softbank Robotics Corp	1	France
Spectronn Inc	1	USA
Stichting Imec Nederland	1	The Netherlands
Understory LLC	1	USA
University of Southern California	1	USA
University of Texas Sys	1	USA
Vadu Inc	1	USA
Verizon Patent & Licensing Inc	1	USA
Vincent, Albert Charles	1	unclear
Wacom Co Ltd	1	Japan
Wilde, Timothy James	1	unclear
XRSpace Co Ltd	1	Taiwan

Received December 2020; revised January 2021; accepted February 2021.