Overview and Commentary of the California Workplace Technology Accountability Act

Airlie Hilliard,[1,2*] Emre Kazim,[1,3] Tom Kemp,[4] Kelvin Bageire[1]

[1] Holistic AI, 18 Soho Square, London, W1D 3QL, UK
[2] Institute of Management Studies, Goldsmiths, University of London, New Cross, London, SE14 6NW, UK
[3] Department of Computer Science, University College London, Gower St., London WC1E 6EA, UK
[4] Kemp Au Ventures, Palo Alto, California, US
[*] Corresponding author: airlie.hilliard@holisticai.com

Abstract

Technological innovation and the resulting automation are increasingly being applied in the workplace in a variety of contexts and while this can remove or lessen the burden of tedious and time-consuming tasks, the workplace is a high-risk context where employment-related decisions can majorly impact a worker's life. Consequently, a number of ethical concerns with automated employment decision systems have been raised, with California proposing the *Workplace Technology Accountability Act* to limit the use of electronic monitoring systems and automated decision systems to specific times of day, activities and locations that must be proven as essential job functions and gives workers the right to know, review and correct data held about them by their employer. In this article, we provide a summary and discussion of the key points of the legislation before providing a commentary, where we identify four key themes: i) the creation of boundaries that can contribute to a healthy work-life balance and protect the privacy of workers; ii) how the requirement for impact assessments of automated decision tools and worker information systems reflects the wider movement towards algorithmic assurance; iii) the necessary and potentially problematic requirement to share notices and impact assessment reports with the Labor Agency; and iv) how the proposed legislation might conflict with existing law while not exempting smaller businesses. Our intended readership is those interested in the regulation of automated employment decision tools, algorithmic assurance, and employers and workers in California affected by the proposed legislation.

Keywords: automation; impact assessments; workplace monitoring; algorithm; assurance

1. Introduction

Automation has vast applications across the globe, ranging from agriculture (Gwagwa et al., 2021), to medicine (Kazzazi, 2021), to self-driving cars (Takács et al., 2018). Among the most high-risk applications of automation is the workplace since employment-related decisions can have a major impact on a worker's life and those who are dependent on them. Here, automation and algorithms are often discussed in terms of their use in recruitment, where alternative assessment formats such as algorithmically scored video interviews (Hickman et al., 2021), game- and image-based assessments (e.g., Hilliard, Kazim, Bitsakis, et al., 2022; Palhano et al., 2019) and chatbots (Nawaz & Gomes, 2019) are increasingly being used to automate the screening of candidates. While the most considered aspect of automation in the workplace is initial employment decisions (whether to hire or promote a candidate), with Illinois (820 ILCS 42, 2020) and New York City (Int 1894-2020, 2021) both passing legislation concerning the use of automated selection tools, automation can also be present within the workplace. Indeed, automated systems are increasingly being used across industries to write reports, fulfil orders, and make medical diagnoses (Chui et al., 2015), as well as to track attendance and manage payroll (Mohan Prasad et al., 2019) and monitor performance (Schumacher & Sihn, 2020). As these systems are increasingly being developed and deployed, it is important that there is sufficient governance of their use. This is the aim of the

wider movement towards artificial intelligence (AI) ethics, which studies the psychological, social, and political impact of AI, drawing on philosophical ethics, and calls for greater governance of these technologies (Kazim & Koshiyama, 2021). To this end, California Assembly Member Ash Karla has proposed the Workplace Technology Accountability Act (*AB-1651 Worker Rights: Workplace Technology Accountability Act.*, 2022) to regulate the use of monitoring tools in the workplace. While we are particularly interested in the sections of the legislation that address the use of automated decision tools (algorithms) and the requirement for impact assessments of such systems, the legislation also defines worker data and sets out expectations when dealing with this data, including limiting the use of general electronic monitoring to collect this data.

In this article, we discuss the proposed Workplace Technology Accountability Act, drawing on the perspectives of multiple disciplines including psychology, philosophy, and computer science. We start by providing an outline of the proposed legislation, drawing on other relevant legislation and guidance from both the United States and the United Kingdom. We then comment on some key themes identified in the legislation, including opportunities for and consequences of non-compliance, how the legislation can facilitate boundaries between work and life outside of work, and assurance of technological systems. Our key takeaways are:

- The legislation is a step in the right direction to ensuring that the privacy of workers is maintained and can contribute to maintaining a work-life balance, particularly for those working remotely;
- The requirement for algorithmic and data protection impact assessments reflects the wider movement toward algorithmic assurance, which is likely to lead to the more responsible use of automated tools and data; and
- While the requirement to share documentation and notices with the Labor Agency and other relevant departments is important in ensuring compliance, it could overwhelm these departments unless effective communication and processing mechanisms are established; and
- As the proposed legislation moves forward, other conflicting legislation already in effect in California and the need for exemptions for smaller businesses and start-ups should be considered.

Our intended readership is those who are interested in how technology might shape the future of work and the governance of automated systems in the workplace, those interested in the broader movement towards AI ethics and assurance of algorithmic and automated systems, and employers and employees in California who want to find out more about the legislation or have similar points of contention.

2. Outline of the Proposed Legislation

In this section, we summarise the proposed legislation. We begin with the overall contributions of the Act before discussing the key definitions and summarising the key requirements of the legislation concerning the collection, use, and storage of employee data, and the use of this data to make decisions, particularly when the data is used by an automated decision system. We discuss the legislation in terms of the impact it will have on both employers and vendors, as well as workers in the state of California.

2.1 Contributions

The legislation asserts that its main contributions include:

> a) The requirement for employers to update their telecommuting plan (required by existing law) to reflect changes in the technology they use;

b) Giving workers the right to know, review, correct or secure data collected about them by their employer and restrictions on how the data can be used; and

c) Guidance on enforcement by the Labor and Workforce Development Agency (shortened to the Labor Agency) and the Department of Fair Employment and Housing.

The Act also requires that the primary responsibility for administration and enforcement is with the Labor Commissioner and that the Department of Fair Employment and Housing should investigate violations in coordination with the Division of Labour Standards Enforcement. Further, the Labor and Workforce Development Agency are required to adopt regulations to enable the administering and enforcement of the legislation, including guidelines on how to manage the coordination of enforcement by the divisions of the Department of Industrial Relations (includes the Division of Occupational Health and Safety and the Division of Workers' Compensation). Further, the Labor Commissioner would be required to contravene a committee of relevant stakeholders, who represent the Department of Industrial Relations and the Department of Fair Employment and Housing, among others.

2.2 Definitions

After providing a summary of the key contributions and amendments that would be required to existing laws, the legislation provides definitions of key terms. In this section, we select some key definitions provided by the legislation and comment on any relevant points of contention. Where we directly quote the Act, we italicise the text and place it within quotation marks.

- **Automated decision system (ADS) or algorithm** – "*computational process, including one derived from machine learning, statistics, or other data processing or artificial intelligence techniques, that makes or assists an employment-related decision*."
    - This definition suggests that the legislation considers the automated decision system and the algorithm as interchangeable. We argue, however, that the two are distinct; an algorithm can be just an aspect of an overall system and there may be multiple algorithms involved in a single system.

- **Data or worker data** – "*any information that identifies, relates to, describes, is reasonably capable of being associated with, or could reasonably be linked, directly or indirectly, with a particular worker, regardless of how the information is collected, inferred, or obtained*."
    - Here, the legislation specifies a number of categories that are covered by this definition, including identifying and biometric information, health-related data, HR, communication, audio-visual, or device usage information, and online information including IP address, as well as contents of the worker's personnel file.

- **Electronic monitoring** – "*the collection of information concerning worker activities or communications by any means other than direct observation, including the use of a computer, telephone, wire, radio, camera, electromagnetic, photoelectronic, or photo-optical system.*"
    - Clearly defining the types of data covered by the legislation, the ways in which employees are monitored or data is collected lessens ambiguity, benefitting both the employer who can avoid unintended non-compliance due to incorrect interpretation, and employees wishing to initiate civil action against their employer for failure to comply.

- **Employer** – "*any person who directly or indirectly, or through an agent or any other person, employs or exercises control over the wages, benefits, other compensation, hours, working conditions, access to work or job opportunities, or other terms or conditions of employment, of any worker*"
    - We note that there is no exemption for smaller businesses and start-ups under the current proposed legislation. This is something that could be problematic for such businesses since the actions required are likely to be costly. Therefore, we call for clarification on whether there will be an exemption from SMBs or at least some leniency, like there is with the New York City legislation, where only employers with over 100 employees must comply. This is something that we endorse across AI regulation (e.g., Kazim et al., 2022).

- **Employment-related decision** – "*any decision made by the employer that affects wages, benefits, other compensation, hours, work schedule, performance evaluation, hiring, discipline, promotion, termination, job content, assignment of work, access to work opportunities, productivity requirements, workplace health and safety, and other terms or conditions of employment*."
    - In defining worker, the legislation notes that this covers anyone providing a service to an employer, whether they are an employee or an independent contractor. Providing this clear definition makes it harder for employers who want to excessively monitor their workers and use this to make a decision to avoid compliance with the legislation by only hiring contractors or workers from agencies.

- **Essential job functions** – the definition provides both examples of objective and subjective sources of data to determine the fundamental duties of a position, favouring objective sources and stipulating that subjective data alone cannot be used to make decisions. Presumably, this is to minimise the influence of potential biases in the reporting of subjective data, which could see the job functions of particular subgroups or positions manipulated to suit the monitoring agenda of the employer.
    - Objective: *"the amount of time workers spend performing each function, the consequences of not requiring individuals to perform the function, the terms of any applicable collective bargaining agreement, workers' past and present work experiences and performance in the position in question, and the employer's reasonable, nondiscriminatory judgment as to which functions are essential."*
    - Subjective: *"Past and current written job descriptions and the employer's reasonable, nondiscriminatory judgment as to which functions are essential may be evidence as to which functions are essential for achieving the purposes of the job, but may not be the sole basis for this determination absent the objective evidence."*

- **"Impact assessment** – "*the ongoing study and evaluation of a data collection system or an automated decision system and its impact on workers.*"
    - Since automated systems can be continuously changing, due to updates in the data available or data collection practices, among other causes, periodic monitoring is often not adequate; a system that was previously compliant or assured may no longer be after changes are made. Therefore, the requirement for continuous assessment of a system is the favoured approach for ensuring that all versions of an algorithm result in minimal harm.

- **Worker Information System (WIS)** – *"a process, automated or not, that involves worker data, including the collection, recording, organization, structuring, storage, alteration, retrieval, consultation, use, sharing, disclosure, dissemination, combination, restriction, erasure, or destruction of worker data. A WIS does not include an ADS."*

    - Here, the proposed legislation makes a distinction between an automated decision system (or algorithm) and a WIS. While distinct, information collected or stored by a WIS may then be fed into an ADS and the outputs of an ADS may also then be stored in a WIS.
    - In addition, the proposed legislation does not provide a distinction between a typical human resources (HR) system and a WIS, which is potentially problematic for employers using HR systems for specific purposes like payroll. We discuss this issue more in section 2.6.

- **Workplace** – *"a location within California at which or from which a worker performs work for an employer."*
    - Based on this definition, employers based in California who employ remote workers outside of the state are not subject to compliance with this legislation.

2.3 Worker Rights

In chapter 2, the proposed legislation outlines requirements when worker data is collected, stored and used by an employer. There are six key themes in this chapter: providing notice, requests for information, accuracy of data, justification for data use, security, and liability of the employer. These very much parallel the privacy rights that consumers have under the CPRA, which we discuss further in section 3.4, with respect to businesses collecting, selling and sharing their personal information, as well as the obligations that businesses have to secure consumers' personal data.

- **Providing notice** – the legislation requires that employers collecting worker data should inform the affected workers that they are doing so before or at the time of data collection. The notice should include the type of data being collected, the purpose for collecting data, how it relates to job function, and how it will be used to make decisions. The only time that employers are permitted to inform workers that data has been collected about them after the fact is if doing so prior to data collection would jeopardise the integrity of an active investigation or would violate other relevant legislation. A copy of the notice given to workers also has to be sent to the Labor Agency.

While the legislation does encourage employers to provide notice of data collection, it also stipulates that providing notice at the time of collection is acceptable. Other legislation passed in the US that requires employers to give notice when using automated tools in the context of recruitment necessitates employers to give much more notice to those affected; the New York City legislation (Int 1894-2020, 2021) requires employers to give ten working days notice before using an algorithmic recruitment tool to collect and process data about an applicant. Giving the affected employees adequate time to digest the notice given to them and do any required research can contribute to workers being able to give more **informed** consent (Hilliard, Kazim, Koshiyama, et al., 2022). However, since ADSs are often black-box systems, the level of informed consent required by these laws is not as stringent as consent required by other data protection laws such as GDPR, which requires unambiguous consent. In contrast, the California Legislation does not specify that workers need to consent to this, just that they need to be informed about the data being collected. It is, therefore, not clear what the obligations are for employers if a worker objects to the data being collected. Since the legislation stipulates that only data relevant to essential job functions can be collected, it is unlikely that employers will reconsider the use of monitoring as they assert that the collection of this data is business necessity. Therefore, this may result in workers who do not wish to be monitored losing their job unless the employer is willing to compromise and provide an accommodation or modified procedure. Notwithstanding this, the requirement to send a copy of the notice to the Labor Agency is likely to reduce the likelihood of non-compliance since employers are less likely to just claim they have given notice to their workers without actually

doing so. Still, there may be some who submit the notice to the Labor Agency without sharing it with their employees since the legislation does not require proof of distribution.

- **Requests for information** – employers that collect, store, analyse, interpret, or share worker data should provide information to workers in an **accessible manner** when they receive a verifiable request.
    - Workers can request information about the types of data an employer has about them, the source of the data, the necessity of the data, how it relates to essential job functions and if it influences any employment-related decisions, if it is used as input for or is the output of an ADS and third-party vendors that collect or receive the data. This should be at no cost to the worker.
    - The information that a worker can request is limited to the information held about them – they cannot request the information of another employee, even on their behalf.

The stipulation that this should be at no cost to the worker is important for creating and maintaining transparency about these monitoring systems. This is particularly true for low-wage workers or those with little disposable income who may otherwise be unable to make such requests if it encumbered them with a financial burden. However, the proposed legislation does not give a comprehensive definition of what it means by accessible manner; the term could simply mean that the information is provided in a way that is easy for workers to access (i.e., not buried somewhere that is hard to find), or that the information must be presented in an assessable way (i.e., concepts are simplified and presented in terms understandable by laypersons). This is likely something that will be clarified as the legislation is debated, with the best-case scenario being that the legislation specifies that information should be both easy to find and in a user-friendly format.

- **Accuracy of Data** – employers should ensure that the data held about workers is accurate and current and workers should have the right to correct inaccuracies if they submit a verifiable request.
    - If an investigation by the employer finds that data is inaccurate, then they should correct the data and inform the worker of this. They should also adjust any decisions or systems that use this data, including those of third parties.
    - If investigations by the employer find the data to be accurate, then the employer should inform the worker that the data is not being changed and the steps taken to verify the accuracy of the data.
    - If the data in question is subjective, then the employer is not obligated to make changes providing that they document that the data is subjective and how it was sourced, and the worker is informed of the decision not to change the data.

Ensuring that workers have access to and can request that the data held about them be updated is an important step towards ensuring that there is greater transparency and accountability. Employers must keep their data up to date since employees can request to see and update it at any time, reducing the likelihood of lazy or clumsy data practices. However, the proposed legislation does not make it clear whether the types of data that workers can access will be restricted. For example, if workers can access their complete personnel file, this could potentially limit the desire to do and effectiveness of 360 reviews and limit how free others feel to speak candidly about a worker or make complaints since it could be easy for the worker to identify who made the comments about them, which could result in worker-to-worker retaliation.

- **Justification for data use** – the legislation stipulates that worker data should only be collected or used if it facilitates essential job functioning, is used to monitor production, assess performance, protect the health and safety of workers, administer wages and benefits, or aid compliance with other relevant laws.

    - The sharing of data is also regulated by the legislation, where certain classes of information, such as health data, are prohibited from being shared with a third party unless required by law. The data is also not required to be shared with the local or state government unless it is needed to provide information, comply with the law, or comply with court-related activities.

- **Security** – employers are required to secure the data they have to the best of their ability and to ensure that there are adequate security measures for the type of data they collect. Security measures can include those that are physical, technical and administrative.
    - In the event of a data breach, the employer is required to provide written notice to the affected employees, including the steps that will be taken to address the impact of the breach. The Labor Agency should also be notified of any breaches.

- **Liability** – if the employer uses a vendor to collect or process data, the vendor must comply with the legislation and the employer is jointly liable if the vendor fails to do so.

Explicitly asserting that the employer is also liable for the actions of the vendor increases the accountability. This is something that was not seen in other US legislations, where the NYC mandatory bias audit legislation and Illinois video interview accountability act fail to mention vendors. This could result in some employers passing the blame for non-compliance on to vendors, allowing them to get around the law. By holding them jointly liable, this encourages employers to check the systems of vendors themselves. This is an approach that is endorsed in the UK in the context of algorithmic recruitment tools, where employers are encouraged to carry out their own checks, if possible, of the systems of vendors before working with them (Recruitment and Employment Confederation, 2021).

- **Violations** – employers incur fines ranging from $5000 to $20000 per violation, except from violations of data security requirements, which come with fines of $100 per affected worker.

Considering that estimates suggest that up to 88% of data breaches are the result of human error or negligence (Tessian, 2020), it is unexpected that fines for employers who experience a data breach are so low since greater penalties are likely to result in employers complying more stringently, therefore implementing additional safeguards to protect worker data. Indeed, in comparison to legislation from other countries, this penalty is much lower; in New Zealand fines can be up to $1300 per person affected by a breech, with a maximum fine of $6300 while penalties issued for data breeches under GDPR are 4% of an employer's annual revenue or 20 million euros, whichever is greatest.

2.4 Electronic Monitoring

In chapter 3, the proposed legislation provides guidance on how to enforce accountability when electronic monitoring systems are used in the workplace. The major themes of this chapter are notice and conditions for monitoring. In this section, we provide an overview of each of these themes, highlighting where the legislation has strengths and weaknesses.

- **Notice** – before employers begin monitoring employees, they must receive notice outlining the purpose of the monitoring, what will be monitored and when, how data will be used in terms of productivity assessment or standards, and disclosure of data with third parties.
    - Employers must choose the least invasive form of monitoring possible for the intended purpose and should inform employees of this and how it is the most suitable form of monitoring if there are alternatives available.

    - Notices informing workers that an employer may use electronic monitoring or reserves the right to is not considered acceptable under the Act – notices should be clear and conspicuous. A copy of the notice should also be sent to the Labor Agency.
    - If there is a significant update to electronic monitoring, then employers are required to notify workers.
    - Employers are required to maintain a current list of electronic monitoring services in use and provide notice to all workers annually (by 1st January each year). They must also provide a copy of this to the Labor Agency by 31st January annually.

The requirement for employers to explicitly inform workers that electronic monitoring will be used, not just that it might, prevents employers from hiding their actions in a cloud of ambiguity. Instead, monitoring is made more transparent, allowing workers to be more informed about the data being collected about them.

- **Conditions for monitoring** – employers are prohibited from electronically monitoring employees unless it will facilitate an essential job function, monitor production, assess performance, ensure legal compliance, maintain health and safety, or administer wages or benefits. Employers must also choose the least invasive form of monitoring possible and use the form of monitoring that is limited to the smallest number of workers and collects as little information as possible to meet the purpose of the monitoring.
    - Monitoring that violates laws or identifies workers exercising their legal rights is banned, as well as monitoring of private areas such as bathrooms or personal residence of workers is prohibited unless strictly necessary for health and safety or security reasons. They should also not have to install applications on personal devices or wear or physically implant monitoring devices unless strictly necessary for job function and limited to certain times and activities.
    - Before using an electronic productivity system, employers must submit a summary of the system to the Labor Agency and the system must be reviewed by the Division of Occupational Safety and Health before implementation.
    - The use of data collected through electronic monitoring should not be used by employers as the sole basis for hiring, termination, disciplinary, or termination decisions; employers must conduct independent assessments to inform decisions and should document whether this additional data corroborates with the data obtained using the electronic monitoring system.
    - Again, the legislation extends liability, with the employer being jointly liable if any vendor they use for electronic monitoring fails to comply.

The limiting of monitoring to specific places, times of day or activity is important for the privacy of workers and ensures that monitoring is only used for necessary activities. This is particularly important for private areas, such as bathrooms, as well as for employees who work remotely at least some of the time. Monitoring of workers outside the workplace can be particularly intrusive, especially if they are working from their own home. The creation of boundaries is important not only for the privacy of workers, but also for creating boundaries and ensuring a work-life balance. We discuss this more in section 3.1.

- **Violations** – fines of between $5000 and $10000 for each violation.

2.5 Algorithms

In chapter 4, the proposed legislation provides guidance on the use of algorithms in the workplace, a term it uses interchangeably with automated decision systems. In this section, we outline the main points of the legislation that are relevant to the use of algorithms or

automated decision systems, and compare ADS and WIS. The major themes in this section of the legislation are notice and decision making.

- **Notice** – employers (or vendors) are required to provide sufficient notice to workers before adopting an ADS and those who are already using one when the legislation comes into force should provide notice to employees within 30 days of the legislation coming into effect.
    - sufficient notices are those that are issued within a reasonable time prior to the introduction of the ADS, are given in the way that routine communications are typically issued, and contains details of the nature and purpose of decisions the ADS will be used to make, the output of the system, data that will be used, and who created and who will run and maintain the ADS. The Labor Agency must also be sent a copy of this notice within 10 days of it being sent to workers.
    - Like with electronic monitoring, the employer must maintain a list of current ADS and share it with workers annually by 1st January, as well as send a copy to the Labor agency by 31st January. Notice should also be given to workers when there are significant changes to the ADS.

As mentioned above, when electronic monitoring systems are being used, employers are only required to provide notice before or at the time of data collection. When algorithms or automated decision systems are being used, however, employees must be notified of this within a reasonable time prior to their use. This is the first time the proposed legislation specifies that workers must be given appropriate notice. However, it does not specify what constitutes an appropriate time. This is something that might be clarified as the proposed legislation is discussed and debated, and might see a requirement similar to the New York City legislation (Int 1894-2020, 2021), where candidates must be informed that automated decision systems are being used at least 10 working days prior to their use, being introduced. However, it is a strength of the legislation, in our opinion, that workers must be notified of the automated decision tools already in use at the time of the legislation coming into effect within 30 days. This is because we could otherwise see employers rush to implement these systems before the legislation comes into effect to avoid having to disclose their use to workers.

- **Decision making** – employers should not use ADS to make employment decisions that violate labor or employment law or to make predictions about behaviour not relevant to job function or the likelihood of workers exercising their legal rights. Facial recognition, gait and emotion recognition and personality predictions are also prohibited, as is the use of customer ratings as input data for an ADS.
    - Before algorithms are used within productivity systems, a summary of the system must be submitted to the Labor Agency and the Division of Occupational Safety and Health must review the system.
    - Employers are prohibited from using ADS outputs about a worker's health to inform any employment-related decision and should not rely solely on the output from an ADS to make any hiring, termination, disciplinary or termination system. Instead, the employer should conduct their own investigation independent of the ADS, including establishing meaningful human oversight by an internal reviewer to corroborate the output using other available data.
        - The reviewer must have the authority, discretion, resources and time to corroborate the output and have sufficient expertise relating to ADS and impact assessments, which may be achieved through education, training or prior experience with similar systems.
        - If the output cannot be corroborated, employers are prohibited from relying on the ADS to make decisions
    - When an ADS is used to make a hiring, promotion, termination or disciplinary decision, workers must be given notice of the decision the ADS was used for,

the data used by the ADS, who created and executed the ADS, a copy of impact assessments (discussed in the section below).
    - Again, employers are jointly responsible for the non-compliance of vendors.

Facial recognition is highly controversial, particularly since there are major disparities in the accuracy of these systems for minority subgroups (Buolamwini, 2018). The ethical concerns raised by their use and the lack of regulation of such tools has prompted calls for greater efforts to ensure that they are more transparent and explainable and their harmful impacts are minimised (Almeida et al., 2021). Presumably as a result of their potential harm, and the fact that facial recognition is unlikely to be needed to monitor the essential job functions of workers, the proposed legislation specifically rules out their use. Further, the legislation rules out the use of customer ratings as input for an ADS, likely because these ratings are subjective and therefore vulnerable to human biases. However, this is impractical for those in customer-facing roles since customer ratings are likely to determine a large proportion of job performance metrics. Therefore, if a worker consistently receives poor customer reviews, this is something that would likely need to be considered when evaluating a worker. We, therefore, call for a reconsideration of this, perhaps by specifying that customer ratings may be included as long as they are not weighted as highly as other, objective measures.

Violations – employers receive a fine of between $2500 and $20000 per violation.

2.6 Impact Assessments

In chapter 5, the legislation outlines the requirement for impact assessments. The type of impact assessment required depends on the system being used. When an algorithm or automated decision system is being used, the legislation requires an algorithmic impact assessment, which must be carried out for each separate position for which the ADS is used to make an employment decision. On the other hand, when a WIS is being used, a data protection impact assessment must be carried out.

- **Algorithmic impact assessments (AIA)** – assessments should evaluate the ADS itself and the development processes used to create the system, including the design and training process or data. The resulting report should include details about the input and output of the system, why an algorithmic approach is superior to a non-automated approach, and an evaluation of risks.
    - Risks should be evaluated in terms of the false positives and negatives that can result from using the system, whether using the ADS violates any legal rights, potential privacy harms, or how the system may negatively impact workers economically or otherwise.
    - If risks are identified in the system, the report should outline the specific measures taken to mitigate these risks, as well as the method used to identify these risks.
- **Data protection impact assessments (DPIA)** – assessments should evaluate the potential for a WIS to have a negative impact on workers and should include a systematic description of the nature, scope, context and purpose of the WIS, as well as the potential risks resulting from the system.
    - Risks include the potential for violation of legal rights, discrimination of protected classes, privacy concerns about invasive or offensive surveillance or potential security breaches, infringement on the dignity and autonomy of workers, and negative economic or other impacts.
    - Where risks are identified, the report should outline the method used to evaluate these risks and recommended mitigation strategies, as well as steps that have already been taken to minimise or eliminate risks.

Both of these assessments must be carried out prior to using the system and retroactively for systems being used before the legislation comes into force and should be continuously updated to reflect any changes to the systems. Assessments should be conducted by an independent assessor with relevant experience, and the assessor should consult with workers personally affected by the systems. As part of this, the assessor is required to make a preliminary assessment available to these workers, who can conduct an anonymous review and provide any necessary feedback. Workers are protected from retaliation from employers. Once an assessment has been conducted, employers should submit the report to the Labor Agency. If health and safety risks are identified, the report must also be submitted to the Division of Occupational Safety and Health and if discrimination or bias is detected, employers should also submit it to the state agency overseeing workplace discrimination. The reports should be written in a precise, transparent, comprehensive and easily accessible way, and should outline the assessment's method, findings, results and conclusions and any resulting modifications made to the system. Workers can anonymously dispute assessments submitted to the Labor Agency and can request an investigation if their employer fails to conduct an impact assessment. Data protection impact assessments required by the Act are therefore arguably more stringent than GDPR, which became the gold standard for data protection, since under this regulation, impact assessments can be carried out by internal parties and then signed off by an impartial and independent data protection officer who must make recommendations without any coercion from the employer.

The distinction that the legislation makes between WIS and ADS is an important one; algorithms can present their own unique challenges so should be assessed in a different way to standard or non-computational information systems. Indeed, some important ethical considerations that result from algorithms cannot be addressed simply by adopting data protection principles, particularly since automated systems can be considerably more opaque than other information systems (Kazim & Koshiyama, 2020b). Further, the use of impact assessments is something that is endorsed elsewhere, including in Canada (Government of Canada, 2021) and the United Kingdom (Kazim et al., 2021; Office, 2022). In section 3.3, we discuss how the requirement for these assessments contributes to greater assurance of the systems. However, the proposed legislation does not make a clear distinction between a regular human resources (HR) system and a WIS. We, therefore, read the legislation as requiring any employer using a HR system as being required to do a data impact assessment, which would be expensive and could therefore see some businesses burdened with this disproportionate responsibility, particularly smaller businesses since, as we have discussed, the legislation does not outline any exemptions for SMEs.

In terms of the data protection risks identified by the legislation, in practice there can be a lack of consensus when identifying these risks. Generally, there is no legislation giving you explicit instructions on how to address these risks, however, supervisory authorities such as the ICO can provide guidance on their website on how to address these risks. We note that the legislation is purposeful with its wording when requiring mitigation strategies to minimise risk instead of aim to completely eliminate them since personal data a highly sensitive and therefore risk associated with its collection, use, and storage can only be completely removed if the data is destroyed and the practice discontinued.

- **Violations** – employers receive a fine of $20,000 per violation.

3. Commentary

In this section, we expand upon our identified key themes in the legislation: the creation of boundaries and how this can help to protect privacy and promote a healthy work-life balance; the requirement to provide a copy of notices and impact assessment reports to the Labor Agency; and how the requirement for impact assessments contributes to greater assurance of systems.

3.1 The creation of boundaries

As mentioned above, the proposed legislation helps to introduce boundaries for worker monitoring, limiting it to only certain times, locations and activities. While this is important for all workers, those in the physical workplace are likely to leave their work at work, i.e., when they leave the site, they are no longer on duty. For those working remotely, particularly those who work from home, the boundaries between home and work can become blurred, and it is not as easy for employers to monitor those who are not physically on-site – they must use software or invasive approaches like requiring webcams to be on at all time to monitor workers. Indeed, during the first wave of the pandemic, many workers were required to work from home. This saw many remote workers working longer hours than they would at the office since it is difficult to create those boundaries when working and living in the same environment, leading to some feeling they are always on the clock (Maurer, 2020). Consequently, workers' work-life balance could become impaired if they find it difficult to switch off, which could result in lower job performance and greater turnover intentions (Fayyazi & Aslani, 2015), impacting both workers and employers.

Further, some employers have been known to use invasive monitoring, tracking the keystrokes of employees, taking pictures of them through their webcams and recording their screens, which can be done without workers knowing (Finnegan, 2020). This can not only invade the privacy of workers, but also their family members or anyone sharing their working space. Therefore, the introduction of clear and unambiguous boundaries and conditions for monitoring employees is important for protecting the privacy of workers both in the workplace and those who work remotely.

3.2 Sharing of reports

The requirement of employers to share copies of notices and impact assessment reports with the Labor Agency is likely to be something that encourages compliance; it is unlikely that employers will fake notices to send to the Labor Agency to appear that they are complying with the proposed legislation. This is particularly since workers have the right to dispute assessments or report their employer for failing to conduct an assessment. The requirement to also share these reports with other, more specialised departments is also a positive since they will be more equipped to deal with discrimination and health and safety issues, mitigating as much risk as possible. However, given that there will likely be tens if not hundreds of thousands or more reports received by the Labor Agency each year, it is questionable how they are going to deal with this. They may become quickly overwhelmed, especially if and when the legislation comes into effect since there will be an influx. This may result in some employers using this to their advantage and not providing the appropriate documentation as they know they are likely to slip through the net. Further, since the reports may need to be considered by multiple departments, this could pose additional issues unless an effective mechanism for communication between departments is established before the legislation comes into effect.

At a federal level, the US is considering the Algorithm Accountability Act (*H.R.6580*, 2022), which will also require the submission of impact assessments. To deal with this, the proposed legislation requires 50 personnel to be hired to evaluate these assessments. A similar approach may be taken in further updates to the proposed legislation, with a dedicated team created to process submitted documentation. This would encourage employers to comply and could result in summaries being published outlining technology use in the workplace each year.

3.3 Towards Assurance

The AI ethics movement has prompted considerations about how algorithms and automated systems can be assured, referring to the need to standardise and operationalise AI ethics principles. Assurance of algorithms and associated systems is comprised of multiple elements, one of which is governance, which can be partly achieved through the use of impact assessments (Kazim & Koshiyama, 2020a). Given that our particular interest in the Workplace Technology Accountability Act is the use of algorithms or automated decision systems, the stipulation that impact assessments must be carried out is a point that we believe merits further discussion. Indeed, the requirement for impact assessments of both ADS and WIS tools will contribute to these systems being assured, which can help to both build trust and identify and mitigate risks associated with them (Barrance et al., 2022; Innovation, 2021; Kazim & Koshiyama, 2020a). This is particularly important for systems used in high-risk contexts like the workplace, where decisions made on the basis of these systems can have a significant impact on an individual's life. This requirement, therefore, reflects the wider movement towards AI ethics and assurance, and is something we would like to see more of across sectors. However, for more comprehensive assurance, other requirements are also needed, including assessments of robustness, transparency, bias, and explainability (Kazim & Koshiyama, 2020a).

3.4 Potential conflicts

While at face value the legislation does signal the greater governance of AI, particularly with respect to the employment context, that there have been calls for, in its current form the proposed legislation conflicts with other laws in California regarding workers' rights over their data. Indeed, the California legislature passed the California Consumer Privacy Act (CCPA) in 2018 (*AB 375*, 2018), which introduced the "employee-employer exemption" that exempts employees from the data subject rights (e.g., the right to be informed, the right to access, the right to delete, etc.) that the CCPA gives consumers (in this case, the employee) as it relates to the consumer's personal information collected by a business (in this case, the employer). The legislation also has an exemption for employers whose annual revenue is not greater than $25 million and or who do not have data on at least 50,000 people, something that the proposed legislation does not do.

The employee-employer exemption was set to expire on January 1, 2021. With the passage of Proposition 24, the California Privacy Rights Act (CPRA) of 2020 (*AB-1490*, 2020), which in effect represents Version 2 of the CCPA, the "employee-employer exemption" was push backed until January 1, 2023 (see Section 1798.145(m) of the CPRA and Section 3(a)(8) of Proposition 24). But the exemption has been relaxed in a few ways with the CPRA: (1) the employer must give the employee a notice of what personal information is being collected; (2) the employer will have the obligations with respect to negligent data breaches (see Section 1798.150) vis a vis an employee's data being breached; and (3) employees can prohibit the sale of their personal information by an employer (see Section 1798.145(n) that references Section 1798.120). The CCPA also In the 2022 legislative session there are two bills (*AB-2891*, 2022; *AB-2871*, 2022) that will make the employer-employee exemption permanent or push it back until January 1, 2026, respectively. Thus, Workplace Technology Accountability Act in effect overrides the "employee-employer exemption" found in California's privacy law.

4. Conclusion

While technology in the workplace can bring about many benefits to both employers and workers, if not used in the right way it can also bring about harm, particularly since the workplace is a high-risk context where decisions can have major impacts on a worker's life. To prevent workplace systems, automated or otherwise, being used with malicious intent, greater governance is needed. The proposed Workplace Technology Accountability Act in

California is, therefore, a step in the right direction to ensuring that there are sufficient boundaries and stipulations in place for collecting, using and storing worker data and using WIS and ADS tools. However, other conflicting legislation and possible exemptions should be kept in mind as this proposal moves forward. In this article, we gave an overview and commentary of the proposed legislation, with our key takeaways being that the proposed legislation could be an effective mechanism through which to protect the privacy of workers and their work-life balance, particularly when working from home and that the required algorithmic and data protection impact assessments signal the wider movement in the field of AI ethics towards the assurance of such systems. However, we also note that there is potential for the required reports and documentation from employers to be mishandled if appropriate systems are not implemented before the proposed legislation comes into effect.

5. References

AB-2891, (2022). https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=202120220AB2891

AB 375, (2018). https://leginfo.legislature.ca.gov/faces/codes_displayText.xhtml?division=3.&part=4.&lawCode=CIV&title=1.81.5

Almeida, D., Shmarko, K., & Lomas, E. (2021). The ethics of facial recognition technologies, surveillance, and accountability in an age of artificial intelligence: a comparative analysis of US, EU, and UK regulatory frameworks. *AI and Ethics*. https://doi.org/10.1007/s43681-021-00077-w

Barrance, E., Kazim, E., Trengove, M., Hilliard, A., Zannone, S., & Koshiyama, A. (2022). Review of the CDEI's Extended Roadmap to an Effective AI Assurance Ecosystem Authors. *SSRN Electronic Journal*. https://ssrn.com/abstract=4081132

Buolamwini, J. (2018). *Gender shades: Intersectional accuracy disparities in commercial gender classification *. 81*, 1–15. http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf

AB-1490, (2020) (testimony of Chau). https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=202120220AB1490

Chui, M., Manyika, J., & Miremadi, M. (2015). Four fundamentals of workplace automation. In *McKinsey Quarterly*. https://roubler.com/sg/wp-content/uploads/sites/49/2016/11/Four-fundamentals-of-workplace-automation.pdf

H.R.6580, (2022) (testimony of Yvette Clarke). https://www.congress.gov/bill/117th-congress/house-bill/6580/text?r=2&s=1

Int 1894-2020, Int 1894-2020 (2021). https://legistar.council.nyc.gov/LegislationDetail.aspx?ID=4344524&GUID=B051915D-A9AC-451E-81F8-6596032FA3F9&Options=Advanced&Search

Fayyazi, M., & Aslani, F. (2015). The Impact of Work-Life Balance on Employees' Job Satisfaction and Turnover Intention; the Moderating Role of Continuance Commitment. *International Letters of Social and Humanistic Sciences*, *51*, 33–41. https://doi.org/10.18052/www.scipress.com/ilshs.51.33

Finnegan, M. (2020). *The New Normal: When work-from-home means the boss is watching*. https://www.computerworld.com/article/3586616/the-new-normal-when-work-from-home-means-the-boss-is-watching.html

Government of Canada. (2021). *Algorithmic Impact Assessment Tool*. https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/algorithmic-impact-assessment.html

Gwagwa, A., Kazim, E., Kachidza, P., Hilliard, A., Siminyu, K., Smith, M., & Shawe-Taylor, J. (2021). Road map for research on responsible artificial intelligence for development (AI4D) in African countries: The case study of agriculture. *Patterns*, *2*(12), 100381. https://doi.org/10.1016/J.PATTER.2021.100381

Hickman, L., Bosch, N., Ng, V., Saef, R., Tay, L., & Woo, S. E. (2021). Automated video interview personality assessments: Reliability, validity, and generalizability investigations. *Journal of Applied Psychology*. https://doi.org/10.1037/apl0000695

Hilliard, A., Kazim, E., Bitsakis, T., & Leutner, F. (2022). Measuring personality through images: Validating a forced-choice image-based assessment of the Big Five personality traits. *Journal of Intelligence*, *10*(1), 12. https://doi.org/10.3390/jintelligence10010012

Hilliard, A., Kazim, E., Koshiyama, A., Zannone, S., Trengove, M., Kingsman, N., & Polle, R. (2022). Regulating the Robots: NYC Mandates Bias Audits for Ai-Driven Employment Decisions. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4083189

820 ILCS 42, (2020). https://www.ilga.gov/legislation/ilcs/ilcs3.asp?ActID=4015&ChapterID=68

Innovation, C. for D. E. and. (2021). *The roadmap to an effective AI assurance ecosystem*. https://www.gov.uk/government/publications/the-roadmap-to-an-effective-ai-assurance-ecosystem

AB-1651 Worker rights: Workplace Technology Accountability Act., (2022) (testimony of Ash Karla). https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=202120220AB1651

Kazim, E., Denny, D. M. T., & Koshiyama, A. (2021). AI auditing and impact assessment: According to the UK information commissioner's office. *AI and Ethics*. https://doi.org/10.1007/s43681-021-00039-2

Kazim, E., Gucluturk, O. G., R. S. Almeida, D., Kerrigan, C., Lomas, E., Koshiyama, A., Hilliard, A., & Trengove, M. (2022). Proposed EU AI Act - Presidency Compromise Text - Select Overview and Comment on the Changes to the Proposed Regulation. *SSRN Electronic Journal*. https://doi.org/10.2139/SSRN.4060220

Kazim, E., & Koshiyama, A. (2020a). AI assurance processes. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3685087

Kazim, E., & Koshiyama, A. (2020b). The interrelation between data and AI ethics in the context of impact assessments. *AI and Ethics*, *0123456789*. https://doi.org/10.1007/s43681-020-00029-w

Kazim, E., & Koshiyama, A. S. (2021). A high-level overview of AI ethics. *Patterns*, *2*(9), 100314. https://doi.org/10.1016/j.patter.2021.100314

Kazzazi, F. (2021). The automation of doctors and machines: A classification for AI in medicine (ADAM framework). *Future Healthcare Journal*, *8*(2), e257–e262. https://doi.org/10.7861/fhj.2020-0189

AB-2871, (2022) (testimony of Low). https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=202120220AB2871

Maurer, R. (2020). *Remote Employees Are Working Longer Than Before*. https://www.shrm.org/hr-today/news/hr-news/pages/remote-employees-are-working-longer-than-before.aspx

Mohan Prasad, K., Sai Nagendra Goru, R., Vamsi, D., & Albert Mayan, M. J. (2019). Automated Payroll Using GPS Tracking and Image Capture. *IOP Conference Series: Materials Science and Engineering*, *590*(1), 012026. https://doi.org/10.1088/1757-899X/590/1/012026

Nawaz, N., & Gomes, A. M. (2019). Artificial intelligence chatbots are new recruiters. *International Journal of Advanced Computer Science and Applications*, *10*(9), 1–5. https://doi.org/10.14569/ijacsa.2019.0100901

Office, I. C. (2022). *Guidance on AI and data protection*. https://ico.org.uk/for-organisations/guide-to-data-protection/key-dp-themes/guidance-on-ai-and-data-protection/

Palhano, D. B., Machado, L. S., & Almeida, A. A. F. (2019). Identifying Player Personality via a Serious Game A Pilot Study Using Item Response Theory. *Proceedings of SBGames*, 575–578.

Recruitment and Employment Confederation. (2021). *Data-driven tools in recruitment guidance*. https://www.rec.uk.com/our-view/research/practical-guides/data-driven-tools-recruitment-guidance

Schumacher, A., & Sihn, W. (2020). Development of a monitoring system for implementation of industrial digitalization and automation using 143 key performance indicators. *Procedia CIRP*, *93*, 1310–1315. https://doi.org/10.1016/j.procir.2020.03.012

Takács, Á., Rudas, I., Bösl, D., & Haidegger, T. (2018). Highly Automated Vehicles and Self-Driving Cars [Industry Tutorial]. *IEEE Robotics and Automation Magazine*, *25*(4), 106–112. https://doi.org/10.1109/MRA.2018.2874301

Tessian. (2020). *Psychology of human error*. https://f.hubspotusercontent20.net/hubfs/1670277/%5BTessian Research%5D The Psychology of Human Error.pdf

Trengove, M., & et al. (n.d.). US Algorithmic Accountability Act overview. *Forthcoming*.