Principles to Promote Responsible Use of AI for Workforce Decisions

By Hodan Omaar  |  August 9, 2021

As the global pandemic has accelerated the digital transformation of the economy, including by forcing many businesses to move to remote work, more employers are accelerating their use of artificial intelligence (AI) to support decision-making about the workforce.

AI-enabled tools can support workforce decisions by helping businesses manage their existing employees, as well as recruit and hire new ones. They can boost productivity among employers, such as by reducing the time needed to hire new employees, increasing retention rates, and improving communications and team dynamics among workers. In addition, these tools may help employers reduce human biases when hiring, decide on compensation, and make other employment-related decisions.

To successfully deploy AI for workforce decisions, employers will need to address potential concerns. These include ensuring that the increased use of AI does not exacerbate existing biases and inequalities, metrics AI tools produce are fair and accurate, increased monitoring of employees is not unduly invasive, and processing of biometric does not reveal sensitive personal information about employees that they may wish to keep private, such as information about their emotions, health, or disabilities.

To address these concerns, several policymakers and advocacy groups have called for new public policies that apply the “precautionary principle” to AI, which says that government should limit the use of new technology until it is proven safe. In short, they favor restricting the use of AI because they believe it is better to be safe than sorry. But these policies do more harm than good because they make it more expensive to develop AI, limit the testing and use of AI, and even ban some applications, thereby reducing productivity, competitiveness, and innovation.
Instead, policymakers should pave the way for widespread adoption of AI in the workplace while building guardrails, where necessary, to limit harms. This report offers eight principles to guide policymakers in their approach to the use of AI for workforce decisions:

1. Make government an early adopter of AI for workforce decisions and share best practices.
2. Ensure data protection laws support the adoption of AI for workforce decisions.
3. Ensure employment nondiscrimination laws apply regardless of whether an organization uses AI.
4. Create rules to safeguard against new privacy risks in workforce data.
5. Address concerns about AI systems for workforce decisions at the national level.
6. Enable the global free flow of employee data.
7. Do not regulate the input of AI systems used for workforce decisions.
8. Focus regulation on employers, not AI vendors.

HOW EMPLOYERS CAN USE AI FOR WORKFORCE DECISIONS

AI is a broad field encompassing an entire class of technologies devoted to creating systems that perform tasks characteristic of human intelligence, such as learning and decision-making. Among others, these technologies include some types of automation, machine learning, recommender systems, and natural language processing (NLP). This report uses the term loosely to cover the breadth of AI technologies organizations use to support workforce decisions, including in acquiring talent, retaining talent, assessing employee engagement, monitoring behavior, and allocating benefits.

Acquiring Talent
The main application area of AI for workforce decisions is talent acquisition, wherein the technology can help employers recruit workers in many different ways. First, AI tools can help attract talent by targeting and personalizing job advertisements. For instance, jobs platform ZipRecruiter has a feature called Candidate Calibration, which suggests job postings to candidates based on other candidates whom employers have already rated highly. The tool uses machine learning—a branch of AI focused on building systems that learn and improve from experience without being explicitly programmed with specific solutions—to learn the types of profiles that would likely suit the position. Another ZipRecruiter tool uses AI to match candidates to jobs based on the information a user shares, such as their location, skills, experience, and interactions with other jobs. For instance,
if a user has interacted with a job advertisement for a sales clerk position in New York City, the platform will suggest similar sales clerk positions in the area.

Second, employers can use AI to sort and review applicants. For example, online platforms such as LinkedIn, TaskRabbit, and Fiverr use ranking algorithms to sort candidates for suitability to each job description. These algorithms draw inferences about a candidate’s suitability based on information from their resumes, such as how many years of experience they have or how proficient they are with specific software. In addition, NLP, a branch of AI concerned with the computational processing of human language, can help review applicants by parsing resumes to extract those with the desired qualifications or experience. Textkernel, for example, can parse resumes in 23 different languages. It also uses semantic search technology, which examines resumes for concepts instead of keywords, meaning employers no longer have to figure out the synonyms, abbreviations, and related terms candidates might put on their resume.4

Third, employers can use AI to support and streamline the recruiting process. According to a 2018 review, 98 percent of Fortune 500 companies already use an applicant tracking system (ATS), a type of recruiting software that collates and enables companies to handle applications electronically.5 Using an ATS, such as those Taleo, Recruitee, and Jobvite Hire offer, has become indispensable for many companies, as these systems can track hundreds of resumes for several job openings with ease and automate routine tasks such as scheduling interviews or following up with candidates. AI services can make these systems even more efficient. For example, when following up with candidates, employers can use chatbots such as the Sense Recruiting Chatbot, a conversational AI assistant that uses NLP to identify information that candidates may need and answer frequently asked questions about job openings and the application process.6 Or, when checking references, AI tools can aggregate and analyze the feedback from respondents.

Fourth, employers can use speech analysis for video interviews of job candidates to help identify their suitability for a particular role. Speech-analysis algorithms assess verbal content, such as word choice, vocabulary, intonation, and inflection. For example, HireVue, a U.S. company that is a leading provider of HR software, used these tools to help screen more than six million videos in 2020.7 The company trains its tools by identifying the competencies required to succeed in a job, which can be specific to an industry or an organization. Then it designs interview questions that allow a candidate to demonstrate these competencies, creates customized performance benchmarks to measure these skills, and runs a population through the assessment to collect data about how they
Machine-learning models then predict the likelihood of success in a job based on this data. Because of cultural and contextual variations in how people express themselves, HireVue trains its models on people from the same culture. For instance, in 2018, HireVue partnered with HR company TalentA to launch a localized Japanese version of its pre-employment assessment models. The company has more than 700 customers using its tools, including multinational organizations such as Unilever, Hilton, GE, and Delta Airlines.

Finally, AI tools can help companies increase diversity in the hiring pool and mitigate unconscious bias in hiring decisions. Consider Textio, an application that analyzes job descriptions for biased, gendered, or off-putting language. Research shows that women apply at a lower rate for jobs in male-dominated fields when the job descriptions use words typically associated with male stereotypes, such as “strong,” “leader,” and “determined.” By swapping words that tend to attract some groups more than others for more neutral words, AI tools can help employers develop job descriptions that expand the pool of candidates that apply. Similarly, AI tools can help employers write more inclusive job descriptions. For instance, there is more gender diversity in the R programming language community than in the Python community, which means data-science job descriptions are more likely to attract women if they include R in the job description. AI-powered blind screening tools can also reduce bias in hiring by filtering out indicators of race, gender, age, and socioeconomic status. For example, New York-based start-up Opus AI’s screening tool integrates with a company’s ATS to automatically redact and anonymize candidates’ identifying information, a taxing manual task, so that recruiters do not use this data when reviewing an application.

Retaining Talent
Using AI and automated tools to retain talent is a growing application area. Employee flight risk models, for instance, can help companies identify the types of employee profiles that are the most likely to leave the company, enabling them to make strategic decisions about whom to target for retention. The algorithms these models use rely on neural networks, which are algorithms that mimic the way the human brain recognizes relationships between different datasets, and data about employees, such as their satisfaction level, promotions in the last five years, latest evaluation, and the number of projects they are working on. For example, Kaggle, a public online community for data scientists, has published two AI models to predict employee turnover with approximately 97 percent accuracy. Similarly, using its Watson supercomputer, IBM has developed an AI model that looks at patterns in data from across the company and predicts which employees are most likely to quit. The algorithms then recommend actions, such as more training or awarding an overdue
promotion, to incentivize them to stay. IBM has cut its attrition rates by 50 percent thanks to its use of AI.  

Similarly, when surveys showed that a team of 700 or so employees at Microsoft were much less satisfied with their work-life balance than their counterparts were, the company used AI to understand why. By examining employee calendars and email usage, the company found workers were spending 27 hours a week on average in large team meetings and significant amounts of time reviewing and replying to emails. It also found that people tended to become more engaged and productive when moving to another department within the company, although many just decided to quit rather than move internally. To improve satisfaction among employees, many of which were engineers with specialized skills who would be hard to replace, Microsoft made it easier to transfer between units and updated its approach to meetings and communication.

Assessing Engagement
Employers can use AI to assess employee engagement and personalize development pathways. Using AI for performance management is perhaps most pervasive in call centers, which have easily measurable metrics for performance and productivity. For instance, employers can use Cogito, a call center AI solution that uses voice and speech recognition software, to get productivity assessments of how employees engage with consumers and provide employees with real-time feedback on their performance. If a call center employee is speaking too quickly, the app will flash an icon of a speedometer; if they sound sleepy, it will display an "energy cue" with a picture of a coffee cup; and if they are not empathetic enough, it will display an icon of a heart. Cogito has solutions for workers in various industries, including health care, insurance, financial services, travel and hospitality, and retail.

Employers can also use tools such as Docebo to better understand how employees are engaging with training and suggest content and courses that better suit their pace and learning method. The tool analyzes metrics on an employee, such as how many courses they have completed and areas they struggled in, and provides personalized recommendations on the suite of courses that would best address their training needs. Companies such as Thomson Reuters, Bloomberg, Uber, Denny's, Cineplex, and Starbucks are already using Docebo to create better training paths for their employees.

In addition, since many public and private organizations have significantly, and likely permanently, increased their use of digital communications and collaboration tools during the COVID-19 pandemic, AI tools can monitor these electronic exchanges to allow managers to understand better how their teams work together, including when some are working remotely.
These tools can perform sentiment analysis on email and text communications to better understand employee morale and address employee concerns. For instance, in 2020, Boston-based start-up Humanyze measured collaboration data from a multinational technology company before, during, and after it transitioned to working remotely in response to the COVID-19 pandemic. The study looks at anonymized data from company emails, calendars, and Slack conversations and finds that only 10 percent of employees worked more than 10 hours a day before the shift to remote work. In contrast, after the change to remote work, 50 percent of employees worked more than 10 hours a day. But, the findings also note that employees were spreading out their work across the day to include more frequent and prolonged breaks. These changes to how employees get their work done can create challenges for coordinating schedules and deter teams from necessary collaboration, leading the company to investigate how it might better facilitate collaboration.

Monitoring Employees

Using AI to monitor employees can help companies improve productivity, protect resources, and safeguard their staff and customers. For example, Uber is one company leading the way in contactless attendance tools to make sure its drivers are who they say they are. In 2020, the company launched Real-Time ID Check in the United Kingdom, which uses facial recognition software to match photos drivers take of themselves to the account holder’s profile picture to ensure the correct driver is using the account and thereby the safety and security of those requesting rides. Automatic time and attendance tools can also benefit office-based companies that use records on employee time to bill their clients and pay their employees correctly. For example, HR company Timerack offers mobile apps and biometric tools that collect and integrate employee data with payroll platforms. The company provides additional AI tools that can analyze employee data, provide detailed insights on specific departments or employees, create alerts for late or missed attendance, and track pay rates. Employers can also use Timerack’s tools in conjunction with a COVID-19 symptom screening platform so employees can self-screen before they enter the workplace, which also helps keep the entire workforce safe.

For other companies, AI tools offer ways to assess and optimize how well employees are performing tasks. Consider pizza chain Domino’s, which recently introduced an AI tool called the Dom Pizza Checker to its Australia and New Zealand locations to ensure employees are making pizzas to the company standard. The tool uses in-store cameras, machine-learning software, and sensors to identify what type of pizza an employee is making, whether the pizza matches the one the customer ordered, and even if the distribution of toppings is correct. If it identifies significant mistakes, the system sends the employee an alert. Or consider Amazon, which uses an AI
system with eye-tracking technologies to monitor how safely its employees drive when out on delivery. The tool collects information about where a user is looking, changes in their pupil size, and whether their eyes are open or closed. If it identifies dangerous behavior such as distracted driving, it suggests actions to the driver, such as taking a break. In tests, the tool reduced accidents by 48 percent and distracted driving by 45 percent.31

Allocating Benefits
Many employers are experimenting with using AI to assess employee compensation, benefits, and wellness. Indeed, a 2018 McKinsey report found that two-thirds of employers are looking to use employee data to improve their performance-management systems.32 Tools such as beqom use AI to draw information about which compensation and incentive models are most impactful in a company and calculates how much compensation an employee should get in a standardized, rules-based way, using various factors such as education, experience, and certifications.33 Using AI to help manage compensation can help employers create more equitable pay across their organizations. It can also provide insights into changing market demands for specific skills, allowing employers to compensate their workers for in-demand skills fairly, forecast future employee costs, and give employees an incentive to develop these valuable skills.

AI tools can also provide health and wellness programs. For example, Johnson & Johnson has developed several systems that leverage AI to help employees improve both their mental and physical health. For instance, the company has developed meQuilibrium, a tool that offers personalized advice on what employees can do to better manage stress, such as tips on practicing mindfulness throughout a busy day.34 Additionally, 90 percent of its employees use the company’s health app, whose AI offers personalized advice on actions they can take during the workday to improve their health, such as lowering their cholesterol.35 Similarly, Amazon uses AI to address the physical pain many of its workers experience while doing their jobs. The company uses automated staffing schedules that rotate employees between jobs that use different muscle-tendon groups, which can help protect workers from musculoskeletal disorders by reducing the time they spend making repetitive motions.36

WHAT ARE THE CHALLENGES TO AI ADOPTION FOR WORKFORCE DECISIONS?
Employers will need to address several concerns in order to successfully deploy the use of AI for workforce decisions. Most critically, they will need to ensure that the increased use of algorithms does not harm workers or exacerbate existing biases and inequalities. Furthermore, they will need to ensure that their use of AI systems to process biometric data does not
unintentionally reveal personal information individuals may wish to keep private from their employer and would violate their autonomy.

**CONCERNS ABOUT AI BIAS**

While using AI offers opportunities to mitigate existing human bias from workforce decisions, employers will also have to make sure algorithms do not, whether intentionally or unintentionally, exacerbate existing biases and inequalities, or introduce new ones. In particular, they need to ensure they have measures in place to detect potential biases in complex and widely used algorithms.

Algorithmic bias refers to the tendency of an algorithm to produce results that have significant differences in accuracy across different groups. Unfair algorithmic bias refers to an algorithm that exhibits differences in accuracy that have deleterious, undesirable outcomes for a particular group. AI systems pose risks for algorithmic bias for three main reasons: They can make more complex decisions, their decisions are scalable, and existing legal oversight may not be sufficient to respond quickly or effectively enough to mitigate potential risks.

First, while some algorithms are relatively simple, such as those used in ATSs to sort and filter job applicants, others can be complex, such as the neural networks used to evaluate candidates’ language and speech in video interviews. These complex algorithms can make very accurate predictions, but in many cases, developers cannot precisely explain how their algorithms make decisions and instead can only express the degree of confidence they have in the accuracy of the algorithms’ decisions. The difficulty arises from the fact that while developers or operators can control what data goes into their systems and instruct algorithms on how to weigh different variables, it can be challenging, if not impossible, to program their systems to explain or justify their decisions. As a result, the complexity of algorithmic systems can be problematic because it creates opportunities for bias to inadvertently influence algorithms. For example, the data algorithms train on can be flawed, such as reflecting historical biases, or be incomplete, which developers or operators could fail to account for.

Second, algorithmic decision-making poses a challenge because a single, widely implemented AI system can make many decisions that impact many individuals. For example, a number of companies using the same technology to screen resumes, and this technology erroneously excludes certain qualified candidates—such as someone with a speech impediment—could have a substantial impact on the ability of these candidates to find a job.
Third, existing legal oversight may not be sufficient to respond quickly or effectively enough to mitigate potential risks. For instance, certain applications of algorithms could cause harms without an operator expressly or obviously breaking the law. An online jobs board that uses a targeted advertising algorithm, for example, may not consider age as a direct variable but still use variables that inadvertently serve as proxies, such as a candidate’s graduation year. Therefore, there is a risk it could unfairly discriminate against members of a certain race for job opportunities, perhaps in ways that may not be immediately obvious to the public, regulators, or even the operator.

There are many ways employers can address these concerns. One is for companies to build and disclose robust oversight and accountability measures for AI systems. In particular, employers should strive to provide confidence to employees that they have a process in place to catch unfair or inaccurate decisions, including by allowing employees to challenge decisions regardless of whether a human or an algorithm makes them. Building more worker-facing tools could also help. For example, some companies use AI systems to monitor absences and trigger reviews after a specific number of missed days. But some of these systems are not employee-facing, so if a worker has an authorized absence, such as for medical leave, until reviews are triggered, they have no way of identifying or rectifying systems that make mistakes, which can be stress inducing. Similarly, many companies use automated tools to extract information from resumes, and applicants often cannot determine whether an AI system has parsed their application correctly. By implementing effective ways for individuals to provide oversight and feedback on AI systems’ decisions, employers can build trust that facilitates increased adoption.

More generally, companies should consider how the use of AI aligns with their broader company values and ethics. Companies that do not demonstrate how the use of AI fits within their overall values, such as social responsibility; worker health and safety; fair pay; and diversity, equity, and inclusion, will struggle to convince workers and the public that they use AI responsibly for workforce decisions. Similarly, companies that narrowly focus on using AI responsibly may overlook opportunities to ensure that their non-AI systems and processes achieve their desired outcomes. For example, Johnson & Johnson has successfully infused AI into its efforts to improve employee health because the company has long been committed to creating a healthy workforce. The company’s employee health campaign first started back in 1979 and has developed over the past several decades to adapt to the changing business needs and employee health concerns. Indeed, Johnson & Johnson’s AI-powered health app that now has 90 percent enrollment evolved from a 1995 health risk assessment that only attained 26 percent buy-in from employees.
PRIVACY CONCERNS

Employers will want to address privacy concerns related to the use of AI for workforce decisions. First, they should ensure that the increased monitoring of employee behavior and activity does not become unnecessarily intrusive; and second, they should ensure that the collection and use of biometric data does not inadvertently reveal information employees may wish to keep private, such as about their emotions, health, or disabilities.

One concern frequently raised about the use of AI in the workplace is employee monitoring may become unduly invasive, stemming in part from the fact that workers may not know how or when their employers are using the technology. For instance, the Trades Union Congress (TUC), a national trade union center representing 48 unions across the United Kingdom, published a report in 2020 that finds that 50 percent of U.K. employees believe their companies may be using AI systems they not aware of. Without transparency on what AI systems employers are using (and not using), and how they collect data about their employees, concerns and rumors can grow and fester, damaging the trust workers have in their employers and stalling adoption. Employers have ascended to one of the most trusted institutions in the world, according to a 2019 study by global communications firm Edelman, with workers trusting their employers more so than the government, nongovernmental organizations (NGOs), industry, and media. In the worldwide survey of more than 30,000 people, 75 percent said they trust their employers to “do what is right,” compared with 56 percent of people trusting industry and 48 percent trusting government. This shift in trust presents a unique opportunity for employers to advance innovation and adoption of AI for workforce decisions by building on the trust inherent to the employer-employee relationship.

Companies should prioritize employee awareness in their AI strategies and make information about what AI systems they are using in the workplace, and why, easily accessible and understandable. Highlighting to employees the privacy benefit of using AI could also help. For instance, AI offers opportunities to increase employee privacy by reducing the number of humans who see their personal information. In some countries, such as the United States, it is legal for companies to ask their employees for their vaccination status after the U.S. Equal Employment Opportunity Commission (EEOC) made it explicit with new guidelines in June of 2021. In the same month, Goldman Sachs began requiring all of its U.S. employees to log their COVID-19 vaccination status and date, and the vaccine maker, in the bank’s internal app for employees and noted this information might be shared with their managers for planning purposes. There is a privacy benefit to companies using AI to automate the processing of vaccination information because people are often more comfortable with computers processing their personal data than they are
with humans.\textsuperscript{46} For example, research shows that individuals prefer dealing with remote entities that use computers to process data over “immediately-present people that could judge them.”\textsuperscript{47} Therefore, the growth of AI creates new opportunities to minimize when and how companies access personal data, thereby increasing employee privacy.

A more complex concern is the data AI systems collect can reveal or enable employers to infer information with varying sensitivity levels, which, if misused, risks autonomy violations and discrimination. Consider Amazon’s AI system with eye-tracking technologies discussed earlier. This system monitors the behavior of delivery drivers by tracking their gaze patterns; however, many studies have found that people with autism react differently to stimuli when driving.\textsuperscript{48} Therefore, an employer may infer from eye-tracking AI software which drivers have autism, even though employees may want to keep this information private.

**THE GOVERNMENT’S ROLE AND FLAWED POLICY APPROACHES**

In many instances, self-regulation, market forces, and tort law can positively shape how companies develop and use AI systems for workforce decisions. For example, when Microsoft introduced a new feature to its Microsoft 365 suite that created a productivity score for a company’s employees based on data about their communication, meetings, collaboration, and teamwork, critics quickly labeled the tool as a mechanism for pervasive surveillance. As a result, Microsoft reduced some of the tool’s capabilities, including removing usernames so that companies could not track the productivity of a particular employee. In a statement, Microsoft said,

> Productivity Score produces a score for the organization and was never designed to score individual users. We'll make that clearer in the user interface and improve our privacy disclosures in the product to ensure that IT admins know exactly what we do and don't track … We always strive to get the balance right, but if and when we miss, we will listen carefully and make appropriate adjustments.\textsuperscript{49}

These forces can unfortunately even put an end to some beneficial use cases altogether. Consider HireVue’s recent announcement that it will discontinue facial analysis screening. The company’s chairman and CEO explained, “Over time, the visual components contributed less to the assessment to the point where there was so much public concern to everything related to AI that it wasn’t worth the concern it was causing people.”\textsuperscript{50}
However, because the overwhelming majority of AI applications for workforce decisions benefit the economy and pose modest and not irreversible risks, the role of governments should be to pave the way for widespread innovation while building guardrails, where necessary, to limit harms.51

But many proposed solutions are ineffective, counterproductive, or harmful to innovation because they embrace the precautionary principle, which is the idea that innovations must be proven safe before employers deploy them, and the notion that the government’s role is to be a speed bump—or even roadblock—to technological progress.52 Currently, these policies treat AI for workforce decisions in three ways: too dangerous to allow (i.e., bans specific uses of AI), too dangerous unless proven safe (i.e., prohibits the technology without special approval), and too dangerous without strict regulatory interventions (i.e., requires the technology to jump through unnecessary and costly hoops before operators can use the technology). These policies are misguided because they impose unnecessary costs, limit innovation, and slow adoption due to hyperbolic fears of AI or failures to recognize that existing or more nuanced regulation would address potential issues.53

1. POLICIES THAT TREAT AI AS TOO DANGEROUS TO ALLOW FOR WORKFORCE DECISIONS
The most extreme form of the precautionary principle leads to bans on specific uses of AI for workforce decisions. These policies stem primarily from the fear that some uses of AI, such as those that rely on biometric information, could lead to “intrusive surveillance practices in the workplace,” which the Australian Human Rights Council states have negative impacts on human rights.54 Consequently, many privacy and civil liberty advocates argue employers should not be allowed to use AI tools such as facial recognition and facial analysis that use biometric data at all.

Bans on Facial Recognition Technology
As described earlier, facial recognition technology has several useful applications for workforce decisions, including clocking employees in and out. But various groups and individuals have called for broad bans on facial recognition because, in part, they fear it could lead to pervasive employee surveillance, discriminatory employment practices, and hackers stealing biometric information from employee databases.

Much of this opposition is based on the false belief that systems that use facial recognition are inaccurate. But the National Institute of Standards and Technology (NIST) analyzed nearly 200 facial recognition algorithms and found that, while lower-quality, less-accurate facial recognition algorithms do display bias, the most-accurate algorithms have low false positives/negatives and undetectable differences among demographics.55
Furthermore, many of the most high-profile critiques of facial recognition do not hold up to scrutiny. For instance, the American Civil Liberties Union (ACLU) has repeatedly made the claim that Amazon’s facial recognition service had an error rate of 5 percent when it was used to compare congressional photos to mugshots. In reality, the error rate would have dropped to zero had the ACLU used the recommended confidence threshold of 99 percent.56

Unfortunately, some governments have already begun to implement bans on this technology. For example, King County Council in Washington State passed an ordinance in June 2021 prohibiting the local government from acquiring and using facial recognition technology. The ban means city buildings cannot use facial recognition to limit access to restricted areas to authorized personnel, which makes these buildings less safe for workers.57 San Francisco and Portland (Oregon) have passed similar ordinances, but Baltimore (Maryland) is currently considering the nation’s most restrictive ban. Its proposed ordinance would prevent the city government from acquiring facial recognition technology and ban most commercial uses, thereby cutting off city businesses and workers from a wide range of beneficial applications.58

In the EU, there is a draft law to strictly regulate facial recognition under the Artificial Intelligence Act (AIA). As written, the regulations do not specifically ban AI systems from using biometric data; they only prohibit law enforcement from using AI systems that provide real-time remote biometric identification in publicly accessible spaces (except in specific time-limited public safety scenarios). However, many in the EU do not think this goes far enough. The EU privacy watchdog, the European Data Protection Supervisor (EDPS), released a statement shortly after the AIA was published stating regrets that the commission had not heeded its calls for a moratorium on facial recognition systems:

The EDPS will continue to advocate for a stricter approach to automated recognition in public spaces of human features—such as of faces but also of gait, fingerprints, DNA, voice, keystrokes and other biometric or behavioral signals—whether these are used in a commercial or administrative context, or for law enforcement purposes. A stricter approach is necessary given that remote biometric identification, where AI may contribute to unprecedented developments, presents extremely high risks of deep and non-democratic intrusion into individuals’ private lives.59

No one wants to live or work in a world with “deep and non-democratic intrusion” into their lives, but employers adopting facial recognition in democratic, rule-of-law nations does not equate to such a world. Governments should not overreact to perceived fears and govern facial recognition technologies from a place of fear and demonization. Instead,
they should have a constructive conversation over what harms this technology can cause and tailor regulations to prevent those harms.

Bans on Facial Analysis
Some governments are considering bans on “facial analysis,” which is a similarly named but otherwise very different technology to facial recognition. Unlike facial recognition tools, which compare images of faces in order to estimate their similarities, facial analysis tools predict features about a person such as their age, gender, or emotion based on a photo.60

In January 2021, the Council of Europe, the EU’s leading human rights organization, published guidelines for legislators and policymakers in the EU regarding how they should regulate the processing of biometric data.61 In its guidelines, the council calls for EU countries to impose a strict ban on facial analysis tools that purport to “detect personality traits, inner feelings, mental health or workers’ engagement from face images.”62 It further states, “Linking recognition of affect … to the hiring of staff … may pose risks of great concern, both at the individual and societal levels and should be prohibited.”63 While the Council cannot make binding laws, its judgments create pressure for member states and companies to shift their practices.64 More recently, the EDPS and European Data Protection Board (EDPB) called for a ban on the use of AI to infer people’s emotions, stating,

Deploying remote biometric identification in publicly accessible spaces means the end of anonymity in those places. Applications such as live facial recognition interfere with fundamental rights and freedoms to such an extent that they may call into question the essence of these rights and freedoms. This calls for an immediate application of the precautionary approach. A general ban on the use of facial recognition in publicly accessible areas is the necessary starting point if we want to preserve our freedoms and create a human-centric legal framework for AI.65

Many have found that facial analysis systems are inaccurate, and there are certainly risks to using inaccurate facial analysis to assess how individuals present themselves when applying for a job.66 For instance, an inaccurate facial analysis tool may incorrectly identify a person displaying an even-tempered state in a video interview as disinterested, thereby harming their chances of getting the job. But many are opposed to using facial analysis technology because affect recognition itself, which is the act of inferring a person’s emotional state from their facial movements, is problematic. In a December 2019 report, citing a study that explored the challenges to Inferring emotion from human facial movements, New York University’s AI Institute stated that regulators “should ban the use of affect recognition in
important decisions that impact people’s lives and access to opportunities typically called emotional expressions or facial expression.”

The problem with this argument is it holds AI systems to a standard that simply does not exist for human decisions. Suppose it is harmful to make a hiring decision based on evaluating a candidate’s emotions or traits from their affect. In that case, it should make no difference whether an AI system or a person made that determination. And humans make workforce decisions on this basis all the time. For instance, a study led by Stanford psychology Professor Jeanne Tsai finds that, in general, American employers look for displays of enthusiasm when interviewing because mainstream American culture associates enthusiasm with leadership. Therefore, they are more likely to favor excited candidates over relaxed ones. But for Hong Kong Chinese, calm and even-tempered states make the best impressions. The difference in how candidates show emotion and how employers interpret it can lead to hiring bias. Provisions to ban facial analysis would only be worthwhile if expanded to all hiring decisions based on affect recognition, regardless of whether technology is involved.

2. POLICIES THAT TREAT AI AS TOO DANGEROUS TO BE USED FOR WORKFORCE DECISIONS UNLESS PROVEN SAFE

Some policies treat workforce-related AI systems as “guilty until proven innocent” and require companies to obtain special permission from the government before they can be sold or used. These include auditing AI systems for bias before entering the market or establishing master regulatory bodies to review AI systems before companies can use them. The major problem with this style of regulation is it slows down the pace of innovation, thereby creating unnecessary roadblocks to the development, testing, and use of new technologies.

Premarket Audits for Vendors of AI Systems

Several governments are considering rules requiring companies that sell AI systems to test them rigorously for bias before entering the market. For instance, the New York City Council has proposed a bill that would make it unlawful for companies to sell automated hiring tools that have not been audited for bias in the year before sale, and require every sale to include a free annual bias audit service that provides the results to the purchaser.

Similarly, the European Commission has proposed in the AIA that “high-risk” AI systems be subject to conformity assessments before being deployed and commercialized in the EU’s internal market. In other words, companies providing or using high-risk systems should rigorously test them for safety, fairness, and privacy. The AIA is clear that all systems used for workforce decisions would be considered high risk:
AI systems used in employment, workers [sic] management and access to self-employment, notably for the recruitment and selection of persons, for making decisions on promotion and termination and for task allocation, monitoring or evaluation of persons in work-related contractual relationships, should also be classified as high-risk, since those systems may appreciably impact future career prospects and livelihoods of these persons.71

Premarket audits have several flaws. First, some unfairly target AI systems. For example, the New York City bill only applies to systems that are “governed by statistical theory” or whose parameters include “inferential methodologies, linear regression, neural networks, decision trees, random forests, and other learning algorithms.”72 Therefore, this bill presumably excludes ATSs—which employers have been using since the 1990s—that use less-complex algorithms to aggregate and sort job applicants into databases employers can filter based on set criteria.73 Yet, these systems can cause as much harm, if not more, than the more modern automated systems that use AI to match the best candidate to a job opening can.

Second, these types of policies overlook how technically challenging it is to audit for bias. Vendors can certainly attest that they have not designed their systems to discriminate against candidates based on protected characteristics intentionally, but it is much harder for them to validate that there is no bias with a particular employer or job listing. For example, an employer might indicate a preference for graduates from a specific set of universities, which may skew the demographics of hired candidates. Or a job listing may include unnecessary qualifications that tend to exclude other groups of candidates. Moreover, for vendors to test that their systems do not exclude candidates based on what algorithms identify as their nationality, gender, race, age, or sexuality, they would need data from employers about the legally protected classes to which applicants and employees belong—information many employers do not collect. And even if they do collect this information, responses to these types of questions are always voluntary. So there will always be gaps in the data, which might lead to mistaken conclusions about bias, especially in small datasets.74

Finally, ex ante regulations make it more expensive and time consuming for companies to introduce new AI applications for workforce decisions.75 The combination of higher costs, delays, and risk of losing intellectual property is likely to deter some companies from launching workforce-related AI products and services at all.76

**Oversight Board for Hiring Assessment Tools**

Stemming from fears that AI is inherently dangerous, some have proposed requiring hiring algorithms to gain governmental approval before being permitted to be used by operators. In the United States, 10 senators
signed a letter in December 2020 requesting information about EEOC’s authority and capacity to conduct research and oversight into hiring algorithms, arguing that there needs to be effective oversight of AI hiring tools. One of the main things Congress wants to know is whether EEOC can request access to hiring assessment tools and applicant data from employers or hiring assessment vendors and conduct tests to determine whether the assessment tools may produce disparate impacts.

However, as the Center for Data Innovation explained in *How Policymakers Can Foster Algorithmic Accountability*, master regulatory bodies to oversee all algorithmic decision-making would be largely ineffective.77 If Congress intends to charge EEOC with reviewing all hiring algorithms or establish a new regulatory body to oversee these tools, it should consider the evidence why.

For one, while it may sound reasonable to task regulators with scrutinizing an algorithm for flaws, it is unrealistic to expect that even the most technologically savvy, resource-flush regulators would be capable of reliably glean meaningful information from examining advanced AI systems and their underlying data, particularly at scale.78 Even getting the information it would need to evaluate AI systems would be difficult. Consider the New York City Council case, which passed a law creating a task force to monitor how municipal agencies use algorithms. One of the task force members lamented that “the task force was given no details into how even the simplest of automated decision systems worked ...This undercut the value of the task force, which aimed to escape the theories and generalizations of the ivory tower to examine how these tools were operating in the real world, using the country’s largest city as our test case. Only we never got the data.” Another member described the task force and its report on ways to make algorithms the city uses fairer and more transparent a “waste.”79

For another, some systems may not need to have regulatory oversight in order to perform safely and effectively for workforce decisions. Creating a requirement for a dedicated oversight board to review algorithmic decisions would make some AI systems impractical and inefficient. Organizations should be given the freedom to decide the degree of human oversight necessary based on the context in which they deploy an AI solution.

3. POLICIES THAT TREAT THE USE OF AI FOR WORKFORCE DECISIONS AS DANGEROUS WITHOUT SOME UNNECESSARY RESTRICTIONS

Some countries have introduced policies that set unnecessary restrictions on AI, including how and when operators can use it. These policies prohibit the use of AI for workforce decisions unless it meets specific and excessive
design or use requirements, including requiring companies to conduct impact assessments, obtain expressed consent to use facial recognition, and design AI systems to be explainable. While policymakers create these laws and regulations to protect against harms, the overall impact is reduced adoption and use of AI for workforce decisions.

Impact Assessments
Much like policymakers can use impact assessments to gather evidence about the potential social or economic impact of certain policies, companies can carry out impact assessments on particular algorithms. Impact assessments for AI systems can be ex ante, focusing on prospective analysis, and ex post, focusing on continuous and historical analysis. Impact assessments can be a valuable tool for minimizing harms from AI systems, but many policymakers are rushing to implement laws before thoroughly assessing the potential unintended consequences of their proposals.

The United States has considered mandatory impact assessments, such as with the Algorithmic Accountability Act of 2019. This bill would direct the Federal Trade Commission (FTC) to develop regulations requiring large firms (defined as those with over $50 million in revenue or that have data about a million consumers or consumer devices) to conduct impact assessments for existing and new “high-risk automated decision systems.” The definition of a “high-risk automated decision system” is broad, encompassing many different types of automated systems, including those that pose a “significant risk” to individual data privacy or security or that result in biased or unfair decision-making; those that make decisions that significantly impact consumers using data about “sensitive aspects,” such as work performance and health; those that involve personal data such as race, political and religious beliefs, gender identity and sexual orientation, and genetic information; or those that monitor a large public space. Unfortunately, this bill misses the mark by unfairly targeting only automated high-risk decision-making, rather than all high-risk decision-making, and targeting large companies, when in reality companies of any size can make harmful decisions at scale using AI systems.

California introduced a similar bill in December 2020 called the Automated Decision Systems Accountability Act. This bill would require any California company that sells automated-decision systems to conduct an impact assessment that identifies whether a given system exhibits disparate impact on individuals in protected classes—and every year following the sale, or after any significant modifications to the system, the vendor must conduct another assessment and make the report available to the Department of Financial Protection and Innovation. These impact assessments would evaluate a system in many areas, including the data
minimization practices in place, the duration for which personal information and the outcomes of the system's decision are stored, and the extent to which consumers can object to or correct a system's decision. Companies would be required to address concerns these assessments identify, but unfortunately, they would not be required to disclose these impact assessments.

Companies should be required to publish general information using anonymized data about what they include in their assessments and what their outcomes are. It is undoubtedly essential to ensure personal data about individuals is kept private, but publishing anonymized data would make people more aware of any potential risks of engaging with a particular algorithmic system and create competitive pressure for companies to reduce this risk. It is unlikely that an average individual would review these assessments themselves, but trusted third parties would likely provide the public with easily digestible recommendations.

Opt-In Consent
Some argue that employers should have to obtain affirmative consent from workers before using AI systems. For example, Jenny Yang, a senior fellow at the Urban Institute, argued in a 2020 congressional hearing that the United States should establish a Workers’ Bill of Rights for Algorithmic Decisions, which would, among other things, give workers the right to know how and when their employer is using an AI system by requiring them to obtain their expressed written consent. These rights, she argued, could “build on the GDPR, which creates a more robust individual rights-based approach to data protection that requires companies collecting or processing data of EU residents to comply with strict transparency, accountability, and data minimization requirements.”

This consent method contrasts with opt-out laws, which require organizations to abide by requests from individuals not to use their data. The differences in impact from using opt-in instead of opt-out consent are striking in three main ways. First, research suggests opt-in regimes are more likely to lead to suboptimal data sharing because most users select the default option for many irrational reasons, including how the question of consent is framed. Second, as the Information Technology and Innovation Foundation (ITIF) explained in “The Economics of ‘Opt-Out’ Versus ‘Opt-In’ Privacy Rules,” there are significant transaction costs associated with obtaining affirmative consent in opt-in regimes. Third, opt-in requirements also make it difficult and impractical for companies to use AI to automate many processes because they must develop and offer at scale a manual process for individuals who choose not to opt in to the automated one.
Still, some states have implemented methods of consent for AI systems for workforce decisions. Illinois, for example, passed the Artificial Intelligence Video Interview Act in January 2020, which requires consent to use facial recognition technology in video interviews. This bill requires any employer that uses an AI system to analyze a video recording of an applicant to notify and provide an information sheet to each applicant in writing before an interview takes place. Once recorded and received, an employer may not share an applicant’s videos, except with people whose expertise or technology is necessary to evaluate the applicant’s fitness for the position and, upon request, employers must delete an applicant’s interviews and instruct any other persons who received copies of the applicant video interviews also to delete the videos within 30 days. Maryland is close to enacting a similar law that would make it unlawful to use facial recognition technologies in interviews without each candidate’s consent. Notably, Illinois’s law does not require that companies provide an alternative method of review, which means if an applicant opts out of an AI-based review of their application, an employer could decide not to consider the applicant for the role.

Algorithmic Explainability
Some policymakers fear that AI will make decisions without any accountability, and that decisions will be flawed, including being biased against underrepresented groups. Therefore, they advocate that decisions made by AI systems be explainable, meaning the systems can articulate the rationale for a given result to a query.

A prime example of explainability laws are Articles 13–15 of the EU’s General Data Protection Regulation (GDPR), which require organizations to provide individuals with “meaningful information about the logic involved” in automated decisions. This means firms must explain how an AI system makes decisions that have a significant impact on individuals. While the EU’s guidelines have clarified that these requirements do not necessarily require full disclosure of the algorithms, the information provided should be “sufficiently comprehensive for the data subject to understand the reasons for the decision.” The GDPR requires data controllers in companies using automated decision-making to “find simple ways to tell the data subject about the rationale behind, or the criteria relied in reaching the decision without necessarily always attempting a complex explanation of the algorithms used or disclosure of the full algorithm.”

Like impact assessments, explainability requirements can be a helpful tool to make AI accountable. But, there is a trade-off between how explainable an AI system is and how accurate it is, which means organizations cannot always comply with requirements to explain the logic involved in an algorithmic decision-making process. And even when companies can potentially explain the logic involved, they may not be able to do so in a
As a result, these regulations will force many businesses to not use certain AI systems, especially more sophisticated ones, even though they may be more accurate, safe, and efficient than the alternatives.

WHAT SHOULD GOVERNMENTS DO?

Given the transformative potential of AI for workforce decisions, policy should tilt toward enabling transformation with this technology. The following are eight principles to help policymakers establish policies to encourage the responsible use of AI for workforce decisions.

1. Make government an early adopter of AI for workforce decisions and share best practices.

Governments should promote the growth of AI for workforce decisions. One way is by adopting the technology for their own purposes. By becoming early adopters, national, subnational, and local governments can promote broader adoption of AI for workforce decisions, which will help reduce risks associated with AI and encourage others to adopt and invest in the technology. For example, to mitigate COVID-19-related unemployment, state policymakers in Indiana partnered with AI company Eightfold to launch an AI platform called the Hoosier Network, which matches job seekers to potential roles. Eightfold is also working with the U.S. Department of Labor to build an application for veterans that analyzes their resumes and matches them with skills needed by employers.

Governments should also help set the quality and performance standards of AI technologies used for workforce decisions. For instance, they should fund independent testing of commercial facial recognition systems and support the development of diverse datasets of faces to foster further algorithmic improvements. Doing so would help fill knowledge gaps ranging from the accuracy of different facial recognition tools to the efficacy of facial analysis tools in affect recognition to the potential uses of these technologies in other specific workforce-related applications. Accelerating efforts to address these gaps could inform more effective and inclusive AI solutions, help reduce risk, and support widespread adoption.

2. Ensure data protection laws support the adoption of AI for workforce decisions.

Supporting the growth of AI for workforce decisions requires nations to have a data protection regime that fosters innovation. Even if a country embraces a light-touch approach to AI, strict data protection laws will impact data processing, which lies at the heart of all AI systems, including those used for workforce decisions. For example, the GDPR gives individuals the right not to be subject to a decision based solely on automated processing, which may limit how efficiently and effectively organizations can use AI for workforce decisions.
Japan has taken a light-touch approach to AI regulation, stating a preference for goals-based rather than rules-based regulation. The Ministry of Economy, Trade, and Industry (METI) rightly explained in a 2020 report that the government’s role is to work with businesses to develop nonbinding guidelines and standards that help them achieve their innovation goals—and cemented this view in a 2021 report in which it stated that “legally-binding horizontal requirements for AI systems are deemed unnecessary at the moment.” As such, Japan does not have any laws banning the development of facial recognition systems, which has enabled two of Japan’s leading manufacturers, air conditioning company Daikin and electronics giant NEC, to partner in 2018 to develop a tool that adjusts office temperature when employees are overly warm and drowsy.

However, Japan is the only country in Asia the EU has exchanged joint adequacy findings with, meaning Japan’s data protection laws are roughly comparable to the EU’s GDPR in limiting the collection and use of data to make automated decisions. Japan has recently moved toward aligning its data protection law, the Act on Protection of Personal Information (APPI), even closer with the EU’s GDPR, echoing the EU’s call to limit data transfer to other countries that do not emulate their laws. The strict rules APPI places on how companies can process, store, and share biometric data means employers have to clear considerable regulatory barriers in order to use systems such as the tool NEC and Daikin developed.

Governments should ensure data protection legislation aligns with their AI goals by limiting their impact on AI innovation to the smallest possible degree. This means, among other things, reducing unnecessary regulatory costs and avoiding undermining important uses of data so as not to reduce the supply of innovative technologies and services.

3. Ensure employment nondiscrimination laws apply regardless of whether an organization uses AI.

Most countries have nondiscrimination laws designed to prevent employment discrimination against particular groups of people in protected classes. These existing regulations can also address many of the new risks AI systems pose. However, it is not always clear how and when existing laws and regulations will apply to various AI systems. For instance, the Americans with Disabilities Act (ADA) is a U.S. law that prohibits discrimination on the basis of disability. According to a senior attorney adviser for EEOC, digital interviews do not violate the ADA as long as businesses provide a “reasonable accommodation, if requested, to enable an applicant to use the digital interview format effectively, or must provide another method for conducting an interview.” It might be relatively straightforward to ensure a person with disabilities has access to and can easily use a hiring tool that employs speech analysis, but some people with
speech disabilities may be unwilling to use these systems due to concerns that their disability may lessen their chances of getting a particular job.

In these situations, it is unclear whether businesses would have to provide an alternative method of review, given that they would have made the digital tools accessible. Such uncertainty can both lead to legal repercussions if compliance requirements are not clear from the outset and hold innovation back. That does not mean policymakers need to implement new, AI-specific regulations related to equity, access, and inclusion, but regulators should review and clarify how existing laws apply to these solutions in order to ensure employers comply with the spirit of these laws in the context of AI.

4. Create rules to safeguard against new privacy risks in workforce data.
AI systems that process biometric data for workforce decisions can add convenience, improve security, and increase productivity. But these systems may also disclose personal information employees would prefer to keep private. For instance, an employer may infer that an employee has autism based on data from an AI tool that uses eye-tracking technology, leading to autonomy violations. While existing laws may protect employees from employers using this information to discriminate against them, individuals may still not want employers to have this information.

The current regulatory landscape leaves gaps in protection for these kinds of privacy risks. Existing definitions of personal and biometric data typically do not account for the processing of biometric information beyond purposes of identification. While many privacy advocates propose addressing these gaps by banning employers from collecting and using biometric data entirely, significant restrictions on the use of biometric data would also curtail the use of many beneficial applications.

Policymakers should carefully craft data privacy legislation to generally allow employers to collect and use biometric data, thereby encouraging innovation in the use of AI for workforce decisions while narrowly restricting certain potentially invasive uses without authorization.

5. Address concerns about AI systems for workforce decisions at the national level.
One considerable challenge to creating a national market for AI systems for workforce decisions is businesses are often subject to the jurisdiction of subnational governments, such as states that make their own rules and regulations. When compounded, a company offering AI solutions across the different regions could face restrictions from each territory it operates in, not including national requirements. For example, suppose an AI company in the United States wants to develop a tool to analyze responses to video interviews and support hiring decisions. In that case, they would have to
abide by broad data protection laws in California, Nevada, and Virginia, obtain consent from job candidates in Illinois, be subject to an audit for bias in New York, and navigate additional differing biometric privacy laws in Arkansas, California, Illinois, Texas, Oregon, and Washington.

This thicket of state laws creates unnecessary and unreasonable compliance costs for businesses and threatens the viability of a national market for AI for workforce decisions. A better approach is to address these policy questions at the national level, such as through comprehensive federal data protection legislation that preempts states.103

6. Enable the free flow of employee data.

AI is poised to significantly impact the global economy, adding $15.7 trillion to the gross domestic product (GDP) by 2030.104 Unfortunately, some governments have implemented digital protectionist policies, such as data localization requirements that mandate that data be stored or processed domestically, in a misguided effort to boost domestic AI development. Many data localization policies target personal, financial, and other data related to HR management and global workforce decisions. While local tech companies may be grateful for the competitive advantage these policies give them by shielding them from foreign competition, the effect is quite different and much less welcome for domestic firms outside the tech sector. These policies isolate them from the international marketplace, limit their horizons for growth, and cut off access to some of the world’s most innovative products and services (as not every service provider will set up local data facilities in every market).105

For example, many of the best-known companies providing AI products and services for workforce decisions are U.S. companies. HireVue, for instance, is a Utah-based AI hiring platform that more than 700 companies use, including Goldman Sachs, Oracle, PwC, Unilever, and Vodafone. Data localization rules such as those in the European Commission’s proposed Data Governance Act (DGA) disincentivize foreign AI companies from introducing their workforce-related products in the EU market because of the additional costs they would incur from having to build out physical infrastructure in every jurisdiction in which they operate. Similarly, restrictions on cross-border data flows in the EU’s GDPR, especially after the Schrems II decision, prevent many businesses from transferring employee or applicant data outside the EU, which limits their ability to use AI tools for workforce decisions.106

In today’s global economy, it is also common for businesses to process data from employees and job candidates outside a company’s home country. There are tens of thousands of multinational corporations around the world that hire hundreds of millions of workers. U.S. multinationals alone employed 43 million workers worldwide in 2018. In a 2021 survey of
500 HR professionals and hiring managers by global immigration company Envoy, 96 percent of respondents said sourcing foreign professionals is essential to their company talent acquisition strategy. Additionally, a 2019 report from U.S. freelancing platform Upwork finds that 38 percent of freelance jobs for the most in-demand skills in the U.S. market came from foreign companies. Protectionist policies that restrict cross-border data flows make it more difficult for these companies to hire internationally and manage their global workforces.

Countries should avoid erecting barriers to restrict the flow of employee data in the global economy. The competition for talent is global—and it is fierce—so if governments want to help their firms succeed, they need to ensure it is as easy as possible to both hire the best and manage them as efficiently as possible. Instead of telling firms where they can store or process employee data, countries should hold employers accountable for managing the data they collect, regardless of where they store or process it.

7. Do not regulate the input of AI systems used for workforce decisions.

Some governments have proposed regulating from where companies can get the data they use to train, validate, and test their AI systems. For instance, the EU Commission has proposed in its white paper on AI that companies only use certain EU-approved datasets that conform to specific EU rules on traceability and data quality for training AI systems. In general, better data leads to more accurate models. But that is not always the case, nor is good data sufficient to ensure accurate or unbiased results. However, these requirements would significantly limit the available data EU companies developing AI systems for workforce decisions could use, and thereby limit the capabilities and performance of these systems. In turn, this would put employees at risk because European data is neither representative nor diverse enough to be used to develop workforce-related systems deployed globally. Moreover, companies having to retrain their AI systems to operate in the EU would introduce additional costs that would be passed on to European employers wishing to procure these systems, thereby reducing their incentive to adopt AI for workforce decisions. Lastly, such requirements would likely exclude many foreign companies from the European market, which would reduce both competition and options for employers. Countries that wish to both see the rapid growth of AI for workforce decisions and ensure that these systems have sufficient, representative data to perform accurately should avoid regulating the data sources these AI systems use.

8. Focus regulation on employers, not AI vendors.

Concern about the potential for unfair algorithms often leads to calls to regulate the underlying technology, such as proposals to regulate AI
systems used for workforce decisions. However, policymakers should keep their focus on employers that operate these systems rather than the vendors that develop these tools because employers make the most important decisions about how their systems impact workers. Moreover, employers are best suited to ensure that the AI systems they use operate as intended, and identify and rectify harmful outcomes.111

In addition, focusing narrowly on providing more oversight of employers that use AI for harmful workforce decisions, rather than on providing oversight of employers who make harmful workforce decisions regardless of the technology involved, would do little to address the bigger problem, which is that some employers exhibit biases in their workforce decisions. Indeed, unless employers with discriminatory workforce practices make changes, even seemingly debiased algorithms will not eliminate this problem.

For example, a 2020 study finds that while debiased ranking algorithms—which employers use to automatically rank applicants for a job—can increase the selection rates of underrepresented candidates, their effectiveness is limited in job contexts where employers have a persistent bias toward a specific gender.112 Researchers in the study asked recruiters to choose 4 candidates from a ranked list of 10 that were generated by different ranking algorithms. They also generated additional versions of each ranked list in which all the data was the same but the candidates' genders were switched from male to female and vice versa. Overall, they found that recruiters for moving assistants, who typically favor workers who are men, consistently hired more men in the study; and recruiters for event staffing, who tend to favor women workers, consistently chose more women.

Focusing antidiscrimination rules and enforcement on employers that use AI systems would be a more efficient way of encouraging vendors to develop responsible AI systems. Doing so would send a market signal to developers about what customers will expect of an algorithmic system and encourage developers to provide algorithms with the necessary capabilities through mechanisms such as transparency, explainability, confidence measures, and procedural regularity, or risk losing market share to competitors that do.113

CONCLUSION
These eight policy principles serve as a blueprint for policies that promote AI adoption for workforce decisions, which could help countries reduce unemployment and bolster the COVID-19-impacted economy. A policy framework built around these principles should maximize the benefits of AI while also minimizing the harms these tools pose. But given the dominant narrative around AI is one of fear, policymakers will need to actively
support the growth of AI for workforce decisions in order to ensure it happens promptly, in part through innovation-friendly regulations.
REFERENCES


8. Ibid.


20. Ibid.


28. Ibid.


38. Ibid

39. Ibid.

40. Ibid.


52. Ibid.

53. Ibid.


Ibid.

Ibid.


Daniel Castro and Michael McLaughlin, “Ten Ways the Precautionary Principle Undermines Progress in Artificial Intelligence.”


Artificial Intelligence Act, (European Commission, April 21, 2021).

A Local Law to amend the administrative code of the city of New York, in relation to the sale of automated employment decision tools, Int. No. 1894.

Hodan Omaar, “NY’s Bill on Automated Hiring Will Dampen Its Recovery Efforts” (Center for Data Innovation, February, 3, 2021),

74. Ibid.
75. Ibid.
78. Ibid.
84. Ibid.
There has been some debate about whether the GDPR establishes a “right to explanation.” While it is not legally binding, Recital 71 states that data subjects should be able “to obtain an explanation of the decision reached.” In addition, the Information Commissioner’s Office (ICO, United Kingdom’s independent data protection authority) emphasized that data processors must be able to ensure individuals can “obtain an explanation of the decision” (See ICO, “Rights Related to Automated Decision-Making, Including Profiling,” https://ico.org.uk/for-organisations/guide-to-the-general-data-protection-regulation-gdpr/individual-rights/rights-related-to-automated-decision-making-including-profiling). This makes it difficult to establish what might be required in the future.


Ibid.


110. Daniel Castro and Eline Chivot, “How the EU Should Revise its AI White Paper Before it is Published.”


ABOUT THE AUTHOR

Hodan Omaar is a policy analyst at the Center for Data Innovation. Previously, she worked as a senior consultant on technology and risk management in London and as a crypto-economist in Berlin. She has an MA in Economics and Mathematics from the University of Edinburgh.

ABOUT THE CENTER FOR DATA INNOVATION

The Center for Data Innovation is the leading global think tank studying the intersection of data, technology, and public policy. With staff in Washington, D.C., and Brussels, the Center formulates and promotes pragmatic public policies designed to maximize the benefits of data-driven innovation in the public and private sectors. It educates policymakers and the public about the opportunities and challenges associated with data, as well as technology trends such as open data, artificial intelligence, and the Internet of Things. The Center is part of the nonprofit, nonpartisan Information Technology and Innovation Foundation (ITIF).

contact: info@datainnovation.org

datainnovation.org