**To: California Department of Technology  
Re: Ethical Challenges of Implementing AI**

**Executive Summary:**  
The growth of AI across various industries have highlighted the importance of considering some of the ethical challenges that come with the implementation of AI-enabled systems. This is something that policy makers and governments should be mindful of especially if governments are implementing automated decision-making tools in social services. Some of these risks and challenges include, but are not limited to fairness, privacy, and accountability. The current status quo has been to have third parties moderate and enforce their own AI policies. However, an analysis of available frameworks, juxtaposed with the actions of various agencies, underlie a lack of risk mitigation when these third parties are left to their own devices. Furthermore, while ethics and conversations around AI are starting to occur in Congress and academia, this was not until a couple of years ago and moreover, the trends are very reactionary. Therefore, it is strongly encouraged that:

1. The Department of Technology establish a statewide AI policy to govern the use of AI prior to implementing any AI-enabled systems in automated decision-making capacities. This would be a policy that would guide the practices of non-government agencies as well.
2. The Department of Technology should establish proactive policies to identify AI risks before they occur. This could include regular audits with humans in the loop to ensure that AI-enabled systems are operating effectively as intended.

*Growth of AI Across Different Industries*  
When looking at the growth rate of AI across different sectors, expectantly we saw AI growing relatively faster in Information and Technology sectors. Another area where there was a sudden uptick in the use of AI was the healthcare and social administration. This, coupled with the growth of Machine Learning tools, can be explained by the fact that more social service departments are integrating automated decision-making tools to assist with backlog, reduce paperwork, and increase efficiency. The features of machine learning allow this technology to be extremely useful in extracting trends, so departments are now relying on Machine Learning to perform predictive analyses and help with the decision-making process. With the ongoing pandemic as governments face increasing demand for public services, many have turned to AI-enabled technologies to process these requests faster and more efficiently.

*Challenges & Current levels of Risk Mitigation*

Given the growth in the use of AI it is vital that governments be cautious of some of the ethical considerations of using AI-enabled tools, especially those that arise in social services. This analysis looked at a wide array of frameworks to discern some of the most common AI challenges cited. The most common AI Challenges cited were ensuring fairness, data privacy, and accountability. Much of the issues relating to fairness have to do with mitigating biases. Since AIs extract patterns from the data that they learn from, if the data collected have biases, so too will the results. For example, the state of Michigan implemented a fraud detection system in 2013 using a biased AI dataset which falsely accused approximately 48,000 individuals of fraudulently collecting unemployment insurance demanding repayment which led to severed wages, homelessness, bankruptcies, evictions, and even suicide. It was later seen that “[93% of the fraud determinations were wrong](https://law.justia.com/cases/federal/appellate-courts/ca6/18-1296/18-1296-2019-01-03.html/).”[[1]](#footnote-1) Governments should be mindful that AI-enabled tools will reinforce any discriminatory practices that are cultivated in the environment in which it is.

As AI progresses, so does its capacity to capture, use, and analyze personal information which can raise serious privacy concerns for government if appropriate safeguards are not in place. Most privacy- related issues surrounding AI are typically centered around data. Mass data collection is used to train AI, but AI systems are also continuously collecting data themselves. A privacy concern is that certain types of personally identifiable information (PII), can unintentionally be collected and stored without the owner’s permission. Lastly, the growth of algorithmic decision making has resulted in an increase in the necessity of accountability for their actions. This is tricky however because there is no current legal frame in which an AI can be held accountable for its own actions. In that case do we hold the developer or the user responsible? Failure to establish clear guidelines related to accountability could undermine trust among both experts and the public which could also limit the benefits of AI.

What was interesting to note in our analyses was that while a majority of frameworks referenced the three risk areas noted above, it is also noteworthy to add the risk areas that do not appear in much of the frameworks but are also of vital importance. This includes accessibility, liberty and awareness of misuse. This, in addition to the fact that while there were 1026 organizations in this study, but at most 26 frameworks discussing these risk categories, illustrate a need for more awareness and guidance in these fields.

The need for guidance is further underscored when analyzing the steps that organizations were currently taking to mitigate these risks. Of a survey of 1028 respondents who represented a wide array of departments from governments to tech companies to other third parties, only 48% had Cybersecurity safeguards, only 30% of respondents admitted to having any kind of privacy policy in their AI use and only 13% had practices in place to mitigate issues of unfairness and inequity (which is noteworthy given that our earlier analysis found this to be the most common risk type in AI use).

*Recommendation # 1: Establishing statewide AI policies prior to the use of AI*

This analysis shows us that AI is becoming more prevalent across different sectors but organizations currently using AI are not taking appropriate measures to mitigate these risks. This could come from a lack of knowledge or, as was seen with the available frameworks, a lack of guidance. Furthermore, there is also no regulatory body ensuring that third parties are complying with AI-risk mitigating policies. Therefore, the government is in the best position to establish a statewide AI policy by which all other agencies are forced to comply with. It is especially of the best interest of governments to do so because most departments do not develop their own AI technologies in house but license them from third parties such as Amazon and Google. However, if these companies do not have the proper protocols in place to mitigate these risks, the AI could act erroneously and have unintended consequences for government departments implementing it. In addition, governments should have these policies implemented prior to the use of AI-enabled technology, specifically in automated decision making in social services. The consequences of overlooking this could include incorrect assessments of individuals and accidentally eliminating benefits and services for those who need them the most.

*Recommendation #2: Taking a proactive stance through regular audits.*

The negative consequences of an AI acting erroneously can have huge implications as we saw with the state of Michigan. Therefore, it is imperative that governments do not rely too heavily on AI tools and have regular audits and humans in the decision-making process to ensure the algorithm is behaving as intended. This is an example of a proactive stance that departments can take before implementing AI-enabled systems. Here proactiveness is emphasized because our analysis showed us that within Congress there is very little mention of AI in legislation until 2017 when AI goes from being mentioned roughly 40 times in total to being mentioned over 150 times across committee reports, congressional research reports and legislation. It is no coincidence that this occurred when it did since the timing of this aligns with the aftermath of the previous elections. This is an indication of how AI policies are only ever discussed in Congress as in a reactionary manner. However, by the time the government reacts to an occurrence, the negative consequences have already occurred necessitating the need to proactively seek out these risks and mitigate them before they have a chance of occurring.

*Conclusion*Mitigation of the risks mentioned above can allow governments to harness the full benefits of AI technology and some of these benefits can be applicable worldwide. AI can be used to achieve the United Nations Sustainability Goals. The data shows that there are over 110 different use cases for AI that can help reach these objectives. For example, there are 29 different AI applications that can increase good health and well-being and 24 different use cases to impose stronger, more peaceful institutions. AI can also be used to achieve gender equality and eradicate poverty. Ultimately, like with many different technologies, there are trade-offs when it comes to the use of AI but if used responsibly, the benefits can be unparalleled.

**Dataset taken from:**

<https://www.kaggle.com/antgoldbloom/the-2019-ai-index-report>

1. Felton, Ryan. "Michigan Unemployment Agency Made 20,000 False Fraud Accusations – Report." The Guardian. December 18, 2016. Accessed September 22, 2020. https://www.theguardian.com/us-news/2016/dec/18/michigan-unemployment-agency-fraud-accusations. [↑](#footnote-ref-1)