AI or EhI: A Game Theoretical Analysis of Canada's Artificial Intelligence and Data Act

Maxwell Keleher
Carleton University
Ottawa, Canada
maxwellkeleher@cmail.carleton.ca

ABSTRACT

ACM Reference Format:
Maxwell Keleher. 2023. AI or EhI: A Game Theoretical Analysis of Canada's Artificial Intelligence and Data Act. In *Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), London, United Kingdom, May 29 – June 2, 2023*, IFAAMAS, 8 pages.

1 INTRODUCTION

Artificial Intelligence (AI) is sure to have a significant impact on everyone's life. A relevant example to anyone at university is the concerns amongst professors that tools such as ChatGPT will allow for more sophisticated cheating in university course work [14]. This example relates to AI as a generative tool, but AI might also be used to make decisions. Amazon experimented with using AI to review candidates for employment, but found that their system was biased against women [6]. In fact, the system would "penalize" candidates whose resume contained "women's" as in "women's chess club captain" as well as candidates who were alumni of women's colleges [6]. AI-powered facial recognition systems seem also seem to have gender biases. Buolamwini and Gebru's research of facial recognition system found that the systems are less accurate at identifying women and men with darker skin than white men [2]. Moreover, the systems are even less accurate at identifying women with darker [2]. These biases become extra problematic when they appear contexts such as policing. Many police forces have begun to use AI tools to identify suspects or allocate resources to certain areas; however, these tools tend to perpetuate biases which harm marginalized communities [7]. While AI can ostensibly provide logical, data-driven decisions, these systems as just as susceptible to bias which can drive harmful outcomes.

Clearly, governments need to step up and regulate artificially intelligent systems to protect the general population from harm. Fortunately, Canada has recently proposed new legislation which includes the Artificial Intelligence and Data Act (AIDA). The purpose of AIDA is to set requirements for those producing artificially intelligent systems to reduce the risk that the systems will harm the public [4]. In this project, I will model AIDA as an extensive form game and look for optimal strategies to understand how we might expect companies and users to act once AIDA is passed into law.

Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), A. Ricci, W. Yeoh, N. Agmon, B. An (eds.), May 29 – June 2, 2023, London, United Kingdom.

2 RELATED WORKS

2.1 Regulations

When discussing digital privacy regulations, there is one piece of legislation that tends to loom over all conversations. The European Union (EU) passed the General Data Protection Regulation (GDPR) [1] in 2018 with the intention of protecting EU citizen's personal data. GDPR affords EU citizens such as the right to access, let people obtain the personal data that companies have about them, and the right to be forgotten, allowing people to have companies delete all personal data about them that the company has collected. It also requires that companies take appropriate measures to protect users personal information and obtain informed consent from its users. Companies found not to be following GDPR can face fines up to 20 million euros or "up to 4% of their total global turnover of the preceding fiscal year, whichever is higher". One key aspect of GDPR is that it is extraterritorial. This means that it applies to any company providing services used by EU citizens, even if they do not operate within the EU.

Canada has written its own legislation to protect Canadian citizens' personal information, the Personal Information Protection and Electronic Documents Act (PIPEDA) [3]. PIPEDA was enacted in 2000; it predates GDPR, but it is not extraterritorial. At a high level, PIPEDA regulates the way that companies are abled to collect and use personal data from their customers. Companies must obtain consent from the people whose data they are collecting, and must state the purpose for which they will use said data. PIPEDA also allows Canadian citizen to access the data which companies have collected and allows them to challenge the accuracy of that information. A study of Canadians understanding of PIPEDA and other privacy regulations found that they, do not to understand how to leverage their privacy rights, even when they are aware that they have privacy rights [18].

My project will focus on the recently proposed Canadian Digital Charter [4]. The Digital Charter seeks to update PIPEDA with the Consumer Privacy Protection Act (CPPA). The CPPA will replace PIPEDA and will provide Canadian citizens with greater control over their privacy. It will make it easier for Canadians to migrate their data, provide a version of the right to be forgotten, and will require that companies obtain informed consent by providing plain language descriptions of how they manage your data. The Artificial Intelligence and Data Act (AIDA) will be the first regulatory framework in Canada which targets artificial intelligence [**?**]. AIDA will regulate the development of artificially intelligent systems, particularly those which pose a risk of harm or bias.

2.2 Game Theoretical Analysis of Regulation

This project is inspired by the work of Zander et al. in their paper "Game-theoretical Model on the GDPR" [16]. They created an extensive form game to model interactions between a company collecting and processing personal data, and a user providing the data. In their model, the turns correspond to the company's decision to follow the requirements of GDPR, and eventually the users decision to read provided consent forms and then actually use the system.

From their model, Zander et al. [16] observe that GDPR generates an asymmetrical game between the company collecting data and the person providing the data. They mention that while their model focuses on a single user providing using the system, they acknowledge that "network effects" could impact users decision to use services. They may, for example, feel social pressure to use social networks which collect personal data. They also question whether GDPR does enough help customers move their data between different service providers or if ambiguity in data ownership results in "lock-in effects" where customers feel that they are must continue to use the services of a company that currently holds their data.

In this project I will follow a similar approach to prepare a Bayesian model of AIDA. Zander et al. [16] did not include any utilities in their model, so I will extend their work by including abstracted utilities based on the results of studies on peoples perceptions of digital privacy and artificial intelligence.

2.3 End-User Perceptions

In my masters thesis [10], I studied end-users privacy perceptions of their computers, phones, and digital assistants. I defined digital assistants as stand alone devices which allow users to interact with artificially intelligent home assistants (e.g Google Home or Amazon Echo). While I was not specifically studying perceptions of artificially intelligent systems, the insights into participants' perceptions of their digital assistants gives an indication of how they might feel about those systems. Curiously, most participants were fairly comfortable with having digital assistants in their home and some doubting that there is any legitimate risk. Some participants seemed to expect that automated systems were more trustworthy than strangers when it came to sharing private information. This aligns with other findings that people tend not to have privacy concerns about their digital assistants [11]. Even when participants do express concerns, they tend not to take any meaningful action to address their concerns [8].

Studies which specifically explore participants attitudes towards artificially intelligent systems tend to find that participants are comfortable with these systems – even when it involves sensitive information. Zhang et al. studied perceptions of artificially intelligent systems in healthcare and found that participants generally had positive perceptions of the systems [17]. Ingrams et al. study of citizens perceptions of AI-decision making in the UK government found that, while they are less trusting of decisions made by AI, participants feel that the systems save time and reduce "red-tape" [9].

In the model I construct as part of this project, I try to represent the level of comfort most of the participants in these studies express towards artificially intelligent systems.

3 BACKGROUND

In this section, I will provide some descriptions of concepts that are critical to understand my project. These descriptions are based off of the lectures from class, and *Multiagent Systems: Algorithmic, Game-theoretic, and Logical Foundations* by Shoham and Leyton-Brown [13].

Imperfect-Information Extensive Form Games. An extensive form game is a way to model interactions between multiple players (which represent actors). It assumes that players will act sequentially as opposed to acting effectively simultaneously and can be represented depicted as a tree of decision nodes. At each node, a given player will choose an action which branches to another decision node (or will end the game). When we assume that player are not aware of all the previous actions, we would call this an imperfect-information game. In an imperfect-information game, a player would need to choose an action at a node without exact knowledge the previous actions and therefore to which node their action will lead. Since costumers typically do not have insight into the operations of companies, I will build my model of AIDA as an imperfect-information extensive form game.

Strategies. A pure strategy is the set actions that a given player would play at each of their decision nodes. They might also adopt a mixed strategy meaning that they would have some probability of playing each action. An optimal strategy is one where the player would not benefit by changing their strategy.

Nash Equilibria. A outcome of a game would be considered a Nash Equilibria when neither player would wish to change their strategy to change the outcome. In this case, both players would be playing their optimal strategy.

4 MODEL

Following a similar approach to Zander et al. [16], I constructed an extensive form game using the Artificial Intelligence and Data Act (AIDA) to set the rules of the game. Also like Zander et al., I have simplified this game to be between just two players. The first player is the entity producing the artificially intelligent system. AIDA [4] defines the entity as follows:

> a person is responsible for an artificial intelligence system, including a high-impact system, if, in the course of international or inter-provincial trade and commerce, they design, develop or make available for use the artificial intelligence system or manage its operation.

I use the symbol P to represent this player. Parts of AIDA only apply to companies which are developing "high impact systems". Since regulators have not yet determined a definition of "high impact system", and to reduce the size of my model, I have decided that P will be a responsible for a system which fits the eventual definition of a "high impact system".

The second player is the end-user who is providing their data for the artificially intelligent system and using the system. They might also be affected by the bias or harms which might come from the system. Unfortunately, AIDA does not explicitly define this player nor does it outline any actions this player might take. Instead, these

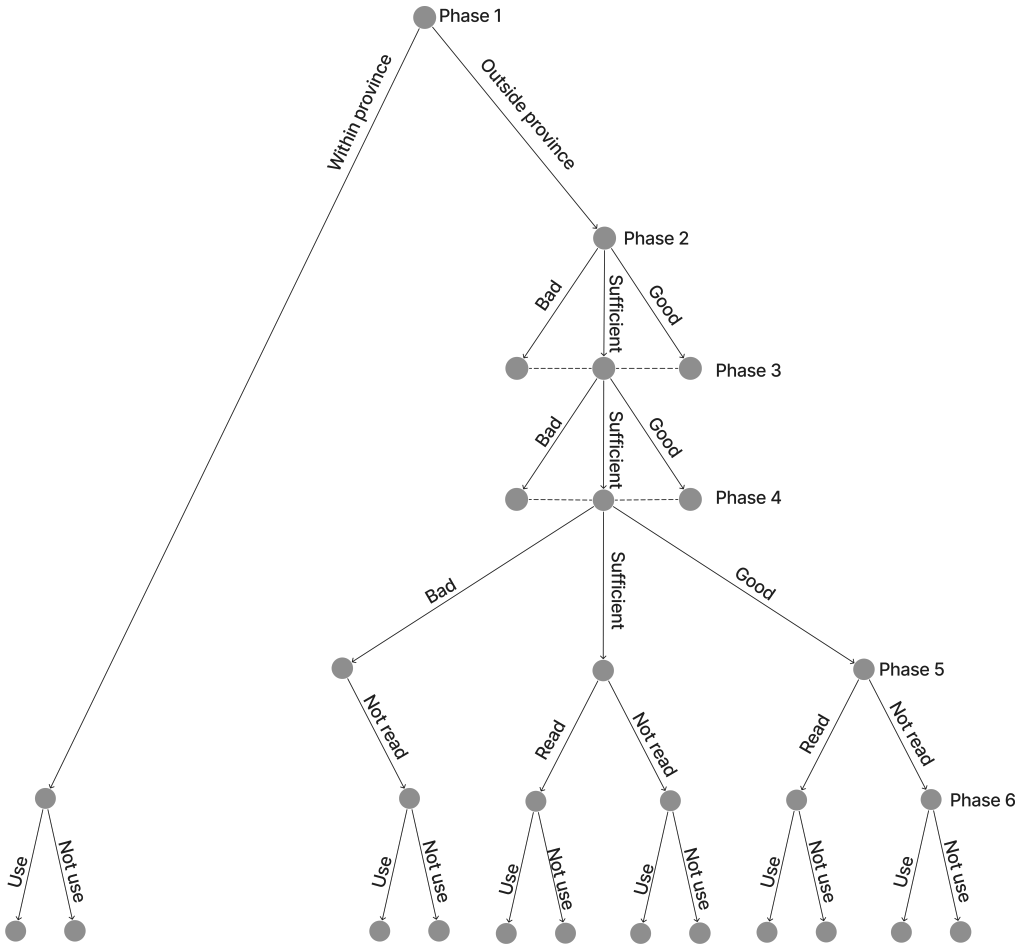

Figure 1: Simplified version of the imperfect-information extensive game I designed to model the Artificial Intelligence and Data Act (AIDA). The company who produces the artificially intelligent system, *P*, will have full information of the game. The user,*U*, will not be aware of *P*'s actions in phases 2 and 3.

actions are implied by the requirements which are placed upon the entities responsible for the artificially intelligent systems. This player is represented by U.

Each players actions depend on the phase of the game. In Figure **??**, I provide a simplified view of my model. For readability, I only show a single branch from the nodes in Phase 2 and 3. U will be unaware of which actions P has taken in Phase 2 and 3 unless they read the description of the system. In the remainder of this section, I describe each of the phases of the extensive form game I have derived from AIDA. In theory, Sections 4.2 and 4.3 could be played in any order, but for the purposes of my model I have set them in the order that they appear in the act itself.

4.1 Phase 1

The language in AIDA which defines the "person responsible" for an artificially intelligent system specifies that the system is used for "international or inter-provincial trade and commerce" [4] Consequently, P could completely avoid AIDA's requirements modeled in Section 4.2, Section 4.3, and Section 4.4 if they restrict their operations to be within a single province. Generally, we would expect that remaining in a single province would provide a slightly negative utility for both the company and the user. Companies would be limiting the growth potential by staying within an province and users from outside the province miss out on their service. It is difficult to image the utilities for both players should P choose to remain within their province. In Section 6, I go into greater detail about when P would take the action to remain within province. The remaining sections of my model focus on P's and U's actions when P has decided to operate outside of their province.

4.2 Phase 2

AIDA [4] specifies that P must take appropriate action to anonymize personal information:

> A person who carries out any regulated activity and who processes or makes available for use anonymized data in the course of that activity must, in accordance

with the regulations, establish measures with respect to
(a) the manner in which data is anonymized; and
(b) the use or management of anonymized data.

For my model, I represent P as having three actions at this phase: Bad, Sufficient, and Good. The Bad action represents P taking insufficient steps to anonymize the data, or failing to anonymize the data at all. Misuse or mismanagement of the anonymized data is also represented by the Bad action. The Sufficient action represents P appropriately anonymizing the data as well as correctly using and managing the anonymized data such that it satisfies AIDA. The Good actions represents P adopting techniques to anonymize, use, and manage their data in a way that surpasses the requirements of AIDA.

By taking the Bad action, P can reduce their cost of doing business (+1 utility), but takes on the risk facing the fines in AIDA if they are audited (see Section 4.7). P can avoid this risk of being fined by taking By taking the Sufficient (0 utility) or Good action (-1 utility due to extra effort). I treat P as having the following preferences: $Bad \succ Sufficient \succ Good$. However, U has an opposite preference given that they would not want to use a systems that unnecessarily requires them to risks revealing their private information. Therefore, U's preference in this phase is: $Good \succ Sufficient \succ Bad$. U gains +1 utility from P taking the Good action, 0 utility from the Sufficient action, and -1 from the Bad action.

4.3 Phase 3

AIDA [4] also specifies that for the companies producing high-impact artificially intelligent systems must appropriately monitor their system, and keep records:

> A person who is responsible for a high-impact system must, in accordance with the regulations, establish measures to monitor compliance with the mitigation measures they are required to establish under section 8 and the effectiveness of those mitigation measures.

Since there are again near infinite ways that a company could act, I have simplified P's actions into three like in Section 4.2. The Bad action represents P keeping insufficient records and/or performing insufficient monitoring of their system. The Sufficient action represents P appropriately keeping records and monitoring the system in a fashion that satisfies the requirements in AIDA The Good action represents P adopting record keeping and monitoring techniques that exceed the demands of AIDA.

Like Section 4.2, P's preferences are $Bad \succ Sufficient \succ Good$ and U's preferences are $Good \succ Sufficient \succ Bad$. These actions have the same impact on both P's and U's utility as described in Section 4.2

4.4 Phase 4

The final component of AIDA which is relevant to my model is the requirement for an entity managing a high impact system to provide a description of their system. Critically, this description is purely informative and does not require that the entity responsible for the system obtain any form of consent from users. AIDA [4] stipulates that:

> A person who makes available for use a high-impact system must, in the time and manner that may be prescribed by regulation, publish on a publicly available website a plain-language description of the system that includes an explanation of
> (a) how the system is intended to be used;
> (b) the types of content that it is intended to generate and the decisions, recommendations or predictions that it is intended to make;
> (c) the mitigation measures established under section 8 in respect of it; and
> (d) any other information that may be prescribed by regulation.

Again, I model P as having three actions available: Bad, Sufficient, and Good. The Bad action represents P preparing a description of their system which does not meet the above requirements. This action also captures failure to provide a description. The Sufficient action represents a case where P provides a description that just satisfies the AIDA requirements. The Good action represent P providing a description of their system which is better than would be necessary for AIDA.

I expect that the description provided in the Good action would be easier for U to read or provide them with unique details. Consequently, the Good action provides U with marginally more utility than the Sufficient action.

Again, P' preferences are $Bad \succ Sufficient \succ Good$ and they have the same impact on utility: +1 for Bad, 0 for Sufficient, and -1 for Good. However, these actions have no direct impact on U's utility; the sufficient description will have greater negative utility if U reads it than the good description.

4.5 Phase 5

For this model, I assume that reading a system description that satisfies AIDA will reveal to U whether P took bad, sufficient, or good action for anonymization (Section 4.2) and for record keeping (Section 4.3). An eye tracking study looking at how people ready terms and conditions found that people do not tend to read terms and conditions unless it is presented to them by default [15]. Since AIDA does not specify that companies present the description at time of use, I assume that U will prefer not to read the published description of system. Therefore, I treat U as having the following preference: $Not\ read \succ read$. I assume that U is able to quickly identify whether P has taken Bad, Sufficient, or Good action in preparing their description without incurring any cost to their utility because it must be made easily accessible. As mentioned in Section 4.4, a good description will have a less negative effect on U's final utility. (-2 for reading a sufficient description and -1 for reading a good description).

I do not believe that P would care whether U has read the provided description; the preference relation for P is: $Not\ read \sim read$

4.6 Phase 6

U's final opportunity to act is when they decide whether or not to use the system. In general, I would assume that people would prefer to use systems which have good anonyimization as well as good record keeping and monitoring. By refusing to use a system,

U receives none of the utility determined by P's actions in the early phases and therefore gets zero utility (not including the negative utility from reading a description). Since U represents the actions of a single user, I assume that their decision about using the system has no impact on P's utility.

4.7 Audit Phase

This phase is different from the other in that neither player takes action and this phase may not occur. In AIDA, the Minister of Innovation, Science, and Technology is afforded the power to audit companies for AIDA compliance.

> If the Minister has reasonable grounds to believe that a person has contravened any of sections 6 to 12 or an order made under section 13 or 14, the Minister may, by order, require that the person
> (a) conduct an audit with respect to the possible contravention; or
> (b) engage the services of an independent auditor to conduct the audit.

AIDA specifies two cases where the Minister could fine an infringing company. The first is anyone who "contravenes any of sections 6 to 12" [4]. These sections correspond to the requirements I describe in Sections 4.2, 4.3, and 4.4. This offense carries fines of up to 10 million CAD or "3% of the [company's] gross global revenue in its financial year before the one in which [they are] sentenced" [4]; whichever is greater. The second offence is more general and applies to anyone using data obtained through illegal means, who recklessly distributes harmful artificially intelligent systems, or who uses an artificially intelligent system to defraud the public [4]. In this case, the offender faces fines of up to 25 million CAD or "5% of the [company's] gross global revenue in its financial year before the one in which [they are] sentenced", which ever is larger [4].

In keeping with the incredibly high limits set for the fines in AIDA, I apply a non-linear structure to determining consequences for P taking the Bad action in multiple phases. P will receive -1 utility if caught taking single Bad action, -10 utility for taking two Bad actions, and -100 for taking only Bad actions. I aim for these punishment values to reflect the fact that these fines will presumably scale with the severity of the offences. I also presume that the failure to comply with a single AIDA requirement could be considered "reckless" behaviours and thus incur the harsher maximum punishment.

Since there is little clarity around what would cause an audit to take place, I use α to represent the possibility that P is audited. In the following section, I will determine the optimal strategies for each player depending on α.

5 RESULTS

In Figures 2, 3, and 4, I provide the utilities for all of the outcomes when P chose to operate outside of a single province. It is unclear how the the choice of action in Phase 1 would affect the utility for both agents, so I focus just on the outcomes when P operates outside of the province. The phrasing of AIDA and my assumptions about the utilities causing in the first player's, P, strategy not to depend on the player who acts after them, U. U's decision to Read or Not read has no impact on P's utility, neither does U's decision to either Use or Not use the system. Therefore, P will choose their actions in the earlier phases such that they maximize their own payoff and effectively ignore the actions of U. U will choose a strategy based on P's optimal strategy.

5.1 *P*'s Optimal Strategies

The only factor which would influence P's strategy, is the threat of facing the consequences in AIDA. As described in Section 4.7, it is unclear what would trigger the audit so the chance that P will be audited is unknown. Instead I hope to reveal what threshold probability of audit would be necessary to prevent P from deciding to ignore the AIDA requirements.

Taking a Good action while taking a Bad action would be irrational as the Good action reduces the utility P would get if they avoid audit without reducing the chance of the audit. The only rational strategy for P to take, which complies with AIDA, is to take Sufficient actions in all phases. This strategy will provide P with 0 utility. There are three groups of strategies which P might take depending on their how many AIDA requirements they are prepared to break. For convenience, I will represent P's strategies as a set with B representing Bad, S representing Sufficient, and G representing Good.

First, they may take a single Bad action while playing Sufficient actions on their other turns: $\{B, S, S\}$, $\{S, B, S\}$, or $\{S, S, B\}$. In this case they would have a $1 - \alpha$ possibility of getting 1 utility, and a α possibility of -1 utility due to the fine. In this case, P will take benefit from this strategy if $\alpha < \frac{1}{2}$ and otherwise should comply with AIDA.

Second, they might take only one Sufficient action: $\{S, B, B\}$, $\{B, S, B\}$, or $\{B, B, S\}$. In this case, P has a $1 - \alpha$ possibility of getting 2 utility, and a α possibility of being fined and receiving -10 utility. This strategy is optimal only when $\alpha < \frac{1}{6}$.

By taking choosing Bad for the anonymization, record keeping and monitoring, and description, P has a $1 - \alpha$ possibility of getting 3 utility, and a α possibility of -100 utility due to the fine. This strategy is only optimal when $\alpha < \frac{3}{103}$.

To ensure compliance with AIDA, and to maximize U's utility, P must believe there to be a greater than 50% chance that they will be audited. This will cause P to take only Sufficient actions.

5.2 *U*'s Optimal Strategies

Now, I explore the strategies U should use to respond to P's strategies. When $\alpha > \frac{1}{2}$, P will adopt the strategy $\{S, S, S\}$ since the odds of an audit are too high. In this case U should avoid reading the description since they will then receive -2 utility whether they use the system or not. By not reading, they receive 0 utility by playing Use or Not use.

When $\frac{1}{2} > \alpha > \frac{1}{6}$, P will play just one Bad action. If P plays $\{S, S, B\}$, U will receive 0 utility whether they use the system or not. If P plays $\{B, S, S\}$ or $\{S, B, S\}$, then U should not read the description, since the final utilities are less than those when not reading. They should then not use the system, since they would receive -1 for playing Use and 0 for Not use. U's best move in this case is to not read the description and to not use the system as they only stand to lose utility by using the system.

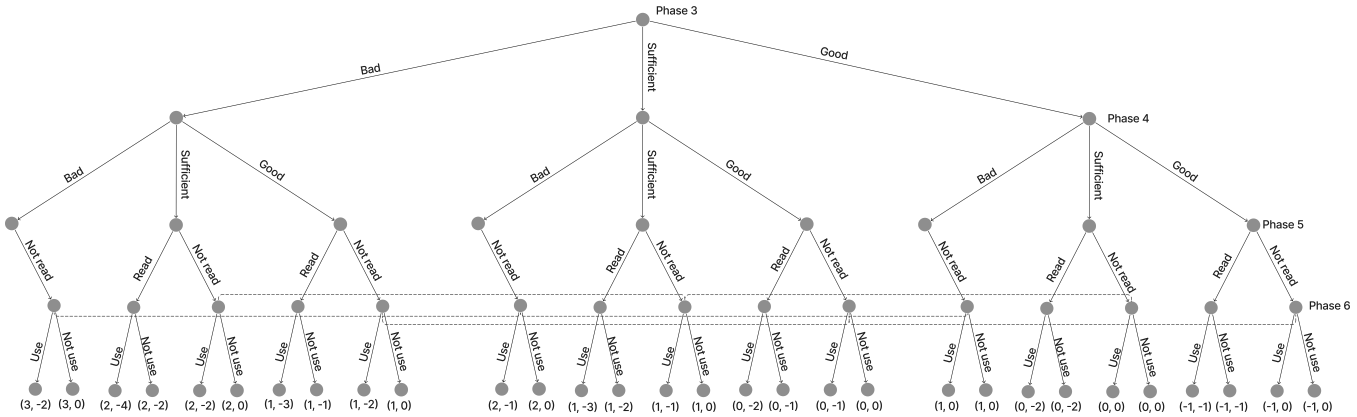

Figure 2: Detailed view of my imperfect information, extensive form game. For legibility, I have focused on the case where *P* plays chooses to act outside the province in Phase 1 and when they take a bad approach to anonymizing the data in Phase 2. Without reading the description, *U* will be unaware of *P*'s actions in Phases 2 or 3. The final utilities for each player appear after *U* acts in Phase 6

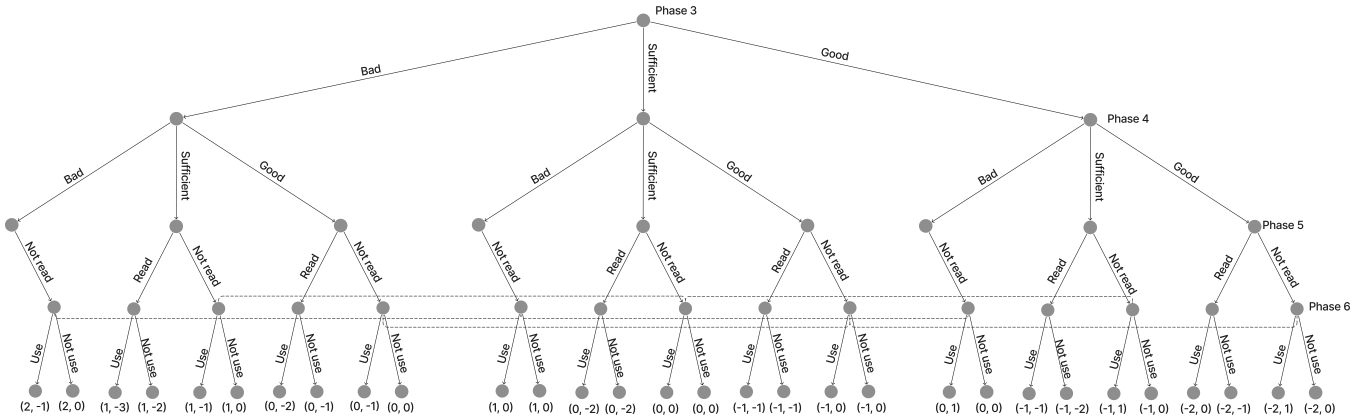

Figure 3: Detailed view of my imperfect information, extensive form game. For legibility, I have focused on the case where *P* plays chooses to act outside the province in Phase 1 and when they take a sufficient approach to anonymizing the data in Phase 2. Without reading the description, *U* will be unaware of *P*'s actions in Phases 2 or 3. The final utilities for each player appear after *U* acts in Phase 6

When $\frac{1}{6} > \alpha > \frac{3}{103}$, P will take two Bad actions. If P plays $\{B, S, B\}$ or $\{S, B, B\}$, then U should not use the system since they get -1 utility for playing Use and 0 for playing Not use. When P plays $\{B, B, S\}$, U best option is to not read and not use (0 utility) if they read. If they were not read the description and use the system, they would get -2 utility which is as good or better than the utilities when they read the description. Again, U is best off when they do not read the description and then do not use the system.

When $\alpha < \frac{3}{103}$, P will play $\{B, B, B\}$. Unsurprisingly, U should not use the system taking 0 utility and avoiding -3 utility from using the system.

Overall, U never seems to benefit from reading the description of the system. U should only feel confident using a system if they think there is at least a 50% chance that infringing companies will be audited. Otherwise they are always better off not using the system.

5.3 Nash Equilibrium

The equilibrium of my game will depend on the the players' perceived likelihood that an audit would catch an infringing company. When $\alpha > \frac{1}{2}$, P will adopt the strategy $\{S, S, S\}$ and U will adopt a mixed strategy of $\{0.5\{Not\ read, Use\}, 0.5\{Not\ read, Notuse\}\}$. When $\frac{1}{2} > \alpha > \frac{1}{6}$, P plays a mixed strategy of $\{0.33\{S, S, B\}$, $0.33\{B, S, S\}$, $0.33\{S, B, S\}\}$ and U will play $\{Not\ read, Not\ use\}$. When $\frac{1}{6} > \alpha > \frac{3}{103}$, P should play $\{0.33\{B, S, B\}$, $0.33\{S, B, B\}$, $0.33\{B, B, S\}\}$ and U should respond $\{Not\ read, Not\ use\}$. Finally, When $\alpha < \frac{3}{103}$, P will play $\{B, B, B\}$ and U will play $\{Not\ read, Not\ use\}$.

6 DISCUSSION

In this section, I will discuss some of the real world problems reflected in my model. While it is a very simplified version of the situation, I think my model highlights some legitimate problems with how digital legislation serve end-users.

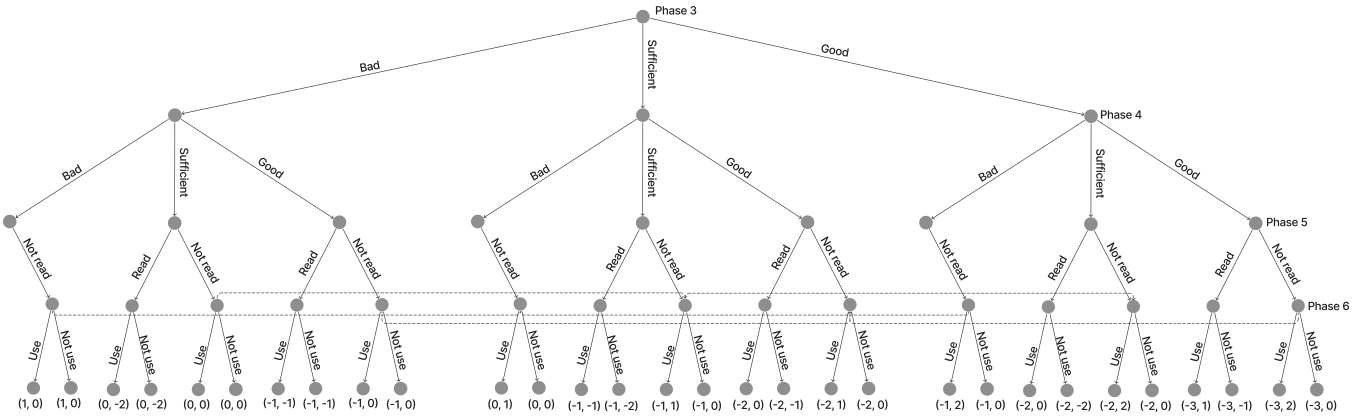

Figure 4: Detailed view of my imperfect information, extensive form game. For legibility, I have focused on the case where *P* plays chooses to act outside the province in Phase 1 and when they take a good approach to anonymizing the data in Phase 2. Without reading the description, *U* will be unaware of *P*'s actions in Phases 2 or 3. The final utilities for each player appear after *U* acts in Phase 6

6.1 Reading the description

In my model, the systems descriptions are essentially useless. A user only benefits from reading a description when the system would cause them to loose greater utility than they would loose from reading the description. I set U's utilities for reading the system description so that it reflected the real world behaviour where people tend not to read privacy policies or terms and conditions. My model certainly reflected this reality. The -2 cost to read a sufficient description meant that users tended to be better off using a partially harmful system with a description they had not read, than if they read the description for non-harmful system.

While this seems absurd, the popularity of social media sites with poor data protection [5, 12] seems to validate my finding that users would rather take the risk of ignorantly using a harmful system, than waste time reading about a system that they do not use. When the company is following AIDA already, reading the description will incur negative utility but the system will provide neutral utility. As long as a user has faith that the company is following AIDA, they seem not to benefit from reading reading the published description. Therefore, the descriptions of systems that companies provide should be focused on explaining the risks that are unique to their system or that a user could not reasonably predict.

There might also be cases where some outside actor pressures a user to use an artificially intelligent system, e.g for work or for school. In this situation the user stands to gain nothing from reading the description. The information in the description can only give users a warning that they might interact with a harmful system, allowing them to avoid the extra negative utility. If they are being forced to use the system anyways, they should avoid reading the system description since it will give them additional negative utility.

It is also worth noting that, since users generally will not read a description, cutting corners when anynomizing the data or monitoring of the system led to worse outcomes for the users. This seems to imply that it would be more beneficial for user if there were harsher punishment for companies that contravene those responsibilities compared to companies who contravene their responsibility to provide a description.

6.2 To use, or not to use (that is the question)

Due to the constraints on the scope of this projects, I assumed that U's, the player who is using the system, choice to use a system would have no direct effect on their utility. In the end, U tended to be equally well off, or even better off, when they chose not to use the systems. I think that this aspect of the model is perhaps least realistic since there are surely benefits that a user might receive from using a system, even if it could be harmful. As I discussed with some of the examples in the introduction, there are also many systems which harm certain people more than others. If the systems are learning from the data of those that use the system, then there are opportunities for biases to snowball. If women tend to find that a system does not serve them well, then they may opt not to use it. This lack of data from women using the system will inevitable cause the system to perform worse for the women who do decide to use the system.

A more complicated model, or a model of a more specific situation, might be able to improve on my own by incorporating some amount of utility that U received when they use a system (even one which presents risks). However, changing the model to account for interest in the systems might require collecting some preference data.

6.3 The Intraprovincial Loophole

Due to the simplicity of the game and lack of clarity for determining utilities, I avoided the part of my model where P decides to operate within a single province. While the utility of constraining business operations within a single province are unclear, the strict requirements of AIDA, and harsh consequences, could cause companies avoid AIDA all together and perhaps collude with other businesses in separate provinces. This could be addressed through provincial and municipal government passing their own versions of AIDA to

address companies that opt to constrain their business so that they can avoid the costs of properly managing the data used in their artificially intelligent systems.

With this gap in legislation, companies which operate in single province might circumvent AIDA by colluding with companies in other provinces. This sort of behaviour definitely appears to be fraudulent, but it does not seem that it violates anything in AIDA. Addressing these collusion behaviours may require cooperation across the departments in the government. Companies which collude internationally present an even greater problem as collaboration between national governments is, typically, far more complicated than collaboration within a nation's government. Even without collusion, there could be opportunities that incentivize companies to remain within province.

Many of the cases where an artificially intelligent system might inflict severe harm could occur within a single province. Canada's medical infrastructure is handled province by province. A company that builds artificially intelligent systems for medical purposes could easily secure contracts with medical service providers that only operate in their province. Errors or biases in such systems could have deadly consequences, but might not fall under the purview of AIDA. Alternatively, a single-province company might focus on serving municipal and provincial police forces. As I mention in the related work, police forces have already began to use artificially intelligent systems. Policing decision have huge impacts on people lives and in extreme situations can lead to unnecessary, and often unjustified, physical harm or death. While it woudl be in

Critically, legislators should not use this "loophole" in AIDA as an excuse to soften the punishments or requirements. While reducing the strictness of AIDA might make it less tempting to try and use the "loophole" companies already stand to gain from minor non-compliance. Instead, it might be best for the federal government to encourage and assist smaller governments with developing this own version of AIDA.

6.4 Limitations and Future Work

AIDA has not yet been passed as a law, so there no exact data which I can reference for my model. I did my best to extrapolate from existing information, but end-users of artificially intelligent systems may have different place different prioritization on companies properly anonymizing their data versus their approach to record keeping and monitoring.

If continuing this project, I would be interested in collecting some subjective data about Canadian citizens' perceptions of artificially intelligent systems. This could allow me to use utility values that better reflect real attitudes and may reveal issues with the legislation I was unable to capture in this project. On the other hand, more realistic utility values could show that made incorrect assumptions and that I pointed out issues with legislation that are unlikely to appear in reality.

I would also be interested in developing a more complicated, computational simulation which could let me include multiple different types of users. This could allow for exploration of how companies behaviours affects users' adoptions of systems which would in turn change companies' payoffs.

7 CONCLUSION

In this project, I have produced a imperfect-information extensive form game model of the proposed Artificial Intelligence and Data Act (AIDA). By analysing my model for optimal strategies, I find that users typically seem best off when they avoid the artificially intelligent systems entirely, and that companies might try to save on resources by not following AIDA unless there are sufficiently high consequences or high likelihoods that they will be caught. Finally, I explore some implications of these finding such as the way that system descriptions might appear redundant to users who already believe that companies will follow AIDA.

REFERENCES

[1] [n.d.]. General Data Protection Regulation (GDPR) – Official Legal Text. https://gdpr-info.eu/

[2] Joy Buolamwini and Timnit Gebru. 2018. Gender shades: Intersectional accuracy disparities in commercial gender classification. In *Conference on fairness, accountability and transparency*. PMLR, 77–91.

[3] Office of the Privacy Commissioner of Canada. 2018. PIPEDA in brief. https://www.priv.gc.ca/en/privacy-topics/privacy-laws-in-canada/the-personal-information-protection-and-electronic-documents-act-pipeda/pipeda_brief/ Last Modified: 2019-05-31.

[4] François-Philippe Champagne. 2022. Digital Charter Implementation Act, 2022. https://www.parl.ca/DocumentViewer/en/44-1/bill/C-27/first-reading

[5] Nicholas Confessore. 2018. Cambridge Analytica and Facebook: The Scandal and the Fallout So Far. *The New York Times* (April 2018). https://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html

[6] Jeffrey Dastin. 2018. Insight - Amazon scraps secret AI recruiting tool that showed bias against women. *Reuters* (Oct. 2018). https://www.reuters.com/article/idUSKCN1MK0AG/

[7] Catherine Halley. 2022. What Happens When Police Use AI to Predict and Prevent Crime? *JSTOR Daily* (Feb. 2022). https://daily.jstor.org/what-happens-when-police-use-ai-to-predict-and-prevent-crime/

[8] Julie M Haney, Susanne M Furman, and Yasemin Acar. 2020. Smart home security and privacy mitigations: Consumer perceptions, practices, and challenges. In *HCI for Cybersecurity, Privacy and Trust: Second International Conference, HCI-CPT 2020, Held as Part of the 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings 22*. Springer, 393–411.

[9] Alex Ingrams, Wesley Kaufmann, and Daan Jacobs. 2022. In AI we trust? Citizen perceptions of AI in government decision making. *Policy & Internet* 14, 2 (2022), 390–409.

[10] Maxwell Richards Keleher. 2023. *Exploring Privacy Implications of Devices as Social Actors*. Ph.D. Dissertation. Carleton University.

[11] Yuting Liao, Jessica Vitak, Priya Kumar, Michael Zimmer, and Katherine Kritikos. 2019. Understanding the role of privacy and trust in intelligent personal assistant adoption. In *Information in Contemporary Society: 14th International Conference, iConference 2019, Washington, DC, USA, March 31–April 3, 2019, Proceedings 14*. Springer, 102–113.

[12] Nick Logan. 2023. If you use TikTok, the app is collecting a staggering amount of information about you | CBC News. *CBC* (March 2023). https://www.cbc.ca/news/canada/tiktok-data-collection-privacy-1.6763626

[13] Yoav Shoham and Kevin Leyton-Brown. 2008. *Multiagent systems: Algorithmic, game-theoretic, and logical foundations*. Cambridge University Press.

[14] Carrie Spector. 2023. What do AI chatbots really mean for students and cheating? *Stanford Graduate School of Education* (Oct. 2023). https://ed.stanford.edu/news/what-do-ai-chatbots-really-mean-students-and-cheating

[15] Nili Steinfeld. 2016. "I agree to the terms and conditions":(How) do users read privacy policies online? An eye-tracking experiment. *Computers in human behavior* 55 (2016), 992–1000.

[16] Tim Zander, Anne Steinbruck, and Pascal Birnstill. 2019. Game-Theoretical Model on the GDPR. *J. Intell. Prop. Info. Tech. & Elec. Com. L.* 10 (2019), 200.

[17] Zhan Zhang, Yegin Genc, Aiwen Xing, Dakuo Wang, Xiangmin Fan, and Daniel Citardi. 2020. Lay individuals' perceptions of artificial intelligence (AI)-empowered healthcare systems. *Proceedings of the Association for Information Science and Technology* 57, 1 (2020), e326.

[18] Leah Zhang-Kennedy and Sonia Chiasson. 2021. " Whether it's moral is a whole other story": Consumer perspectives on privacy regulations and corporate data practices. In *Seventeenth Symposium on Usable Privacy and Security (SOUPS 2021)*. 197–216.