The idea of AI solving almost all our human problems is fantastic. I often joke about how ChatGPT is the god I’ll believe in. However, we must identify a few risks here. I’d like to talk about the two biggest concerns for this company’s plan. The first one and the biggest concern I have is the black box algorithm and its outcome. The second one is the opt-out approach for giving personal data for data training purposes.

First and foremost, the fact that they don’t know how the algorithm arrives at this result is a significant red flag and unacceptable circumstances. Because that means the company doesn’t have control over its service or products. When the public finds out, they will lose trust in them due to the fact that they are not reliable for their own work. It violates the principles of Transparency and openness, Interpretability, and Control according to Business Ethics and Artificial Intelligence[[1]](#footnote-1). In the briefing, it is mentioned that *“To have full control over AI systems, it is important that both companies and algorithm designers only work with technology that they fully understand. Being able to explain the functionalities of a technology of which they appear to be in control of is essential to build trust with employees, customers and all stakeholders.“*

People need to trust AI, especially in this uncertain world. AI companies should make an effort to ensure people are able to trust the company’s AI. If they don’t know why and how the AI came up with the algorithm, it’ll damage the company’s professional reputation. They could lose the business to other competitors who have people’s trust more. For example, Facebook has a bad reputation for how it uses or even exploits its users’ data. If there’s a platform that has the same functionalities and the same amount of traffic (users), I believe most people will no longer use Facebook. I know a couple of people who value their privacy and don’t use Facebook or its related products at all.

HFII should also follow the principle of Fairness. Unfair outcomes leading to discrimination[[2]](#footnote-2), and we don’t want that. I really don’t understand how the algorithm came up with the result in favor of white males, and people who go to music festivals get a lower credit score. Don’t people have good credit can afford festivals and dining out as well? As a former software engineer, I think when the AI produces an unfair result, the company should first do research or an RCA (root cause analysis). They should check whether the source data is biased from the beginning. Perhaps they targeted more white or male users, so their users have a larger demographic representation in this group, which results in a bias in favor of this group of people. As a female and a person of color without a white sounding name, it always triggers me when I read about these kinds of incidents[[3]](#footnote-3)[[4]](#footnote-4). Perhaps the company can scramble its users’ data by hiding names, genders, races, etc, to make these attributes completely anonymous when they enter their data for data training. This way, even if they don’t know how the algorithm works, they can be clear of discrimination.

The second issue I have is the opt-out method for data collection. Existing users will only be informed about this new system and not be provided with any detailed information, such as how it works or what data is collected. Moreover, users will be automatically enrolled in this service if they don’t opt out. New users will be told that the tools are free, but they have to provide their very personal data to the company for data collection. For me, this is not only ethically inappropriate but also violates some privacy principles. As a golden rule believer personally, I think this is tricky and in a grey area in terms of law.

According to PIPEDA’s section 5, Fair Information Principle 5 – Limiting Use, Disclosure, and Retention, HFII should obtain fresh consent if it intends to use or disclose personal information for a new purpose.[[5]](#footnote-5) Here, they don’t really obtain a fresh consent. Informing the user is not consent, especially the data is very sensitive personal and private. There are no indications of where the company is located. Let’s assume it’s in a state where PIPEDA practices, then the company violates it. If it is in Quebec, then an opt-in method is required.[[6]](#footnote-6)

In conclusion, trustworthy AI is crucial for a company to maintain a good reputation, earn people’s trust, and subsequently expand or grow its business. To maintain a sustainable business, HFII should follow at least these ethical AI principles: transparency, fairness, and honesty. Although I often complain about how Europe is so behind modern technologies after living in Berlin for 8 years, I’m very glad that EU took the first step to protect people’s privacy so I have the chance to reject all the cookie requests everytime I visit a website or service, and I see that more countries such as Canada follows.

After seeing how Facebook exploit our data, I’m very glad that EU has super strong and strict rules on data privacy

I really appreciate GDPR. Thanks to it, I have the opportunity to reject all the cookie requests. Although it might be annoying at times to choose at every website I visit, it does make me feel more secure in an age like this.

What kind of risk there could be?

Purchase habits are personal, private and sensitive.

don’t like this kind of tricky behavior from the big corporates.

**“Respect of privacy**”

What in a world would you do such tricky thing? This will face a ton of critiques I’m not sure if it’s even legal in Canada

In some cases,

AI applications have been referred to as a ‘black box’

where not even engineers can decipher why the machine

made a certain decision.[[7]](#footnote-7)

I don’t think this is true, they probably didn’t dig deeper enough, I believe everything happens for a reason.

Altho it might be as complex as human’s neuron system that nobody has a holistic understanding of it but there should have some trace of evidence to indicate how this AI algorithm works in this way

Think about the users demography. Is it because the app is targeting on white mele or reckless spenders

If they don’t know why, they can’t convince people

They should have an explanation so the general public or potential users know they are genuine

HFII will offer this app to people for free – so maybe will only attract people with cheap minds

*we shape our buildings*

*and afterwards our buildings shape us*

After seeing how Facebook exploit our data, I’m very glad that EU has super strong and strict rules on data privacy

Thanks to GDPR I have the opportunity to reject all the cookie requests

don’t do things you don’t like other to do on you

As a minority, and a golden rule believer when I saw

The company will have to be liable for the information provided to other loan companies (<https://www.torkin.com/insights/publication/bc-tribunal-confirms-companies-remain-liable-for-ai-chatbot-created-information>) companies remain liable for the actions of their AI tools.

Have they given consent for this purpose

Openness principle 8

Unfortunately theres no information on where the company is, because there are different laws regarding the privacy laws regarding data base breach

I have to say white especially male privilege does exist as many evidences show.

For example, white last names will get more interview invites, or male white names get more respect on email exchanges.

I wonder how much “tweaks” the company should involve in DEI policy

As a former software engineer, I suggest the company to do a RCA (root cause analysis) and try to find out why and how did this happen. By this I mean the fact that the algorithm came up with the result in favor of white male, and people who go to music feastivals get lower score. I really can’t scratch my head off how this could happen. Maybe the data they collected is bias from the beginning? Maybe because the users are more whites and male than people of color and female, and that’s why it’s mentioned that people of color are more prone to error.

They basically don’t understand anything and want to start implement this feature to the users?

I would like to suggest them to check if the source of data is bias? If the users are more male than female? More white men than people of colors? Do they have enough data for sampling?

In conclusion, trustworthy AI is very important for a company to have good reputation to earn people’s trust then expand or grow their business. It’s crucial to have a sustainable business

1. Institute of Business Ethics. Business Ethics Briefing Issue 58. Business Ethics and Artificial Intelligence [↑](#footnote-ref-1)
2. Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. Big Data & Society, 3(2). https://doi.org/10.1177/2053951716679679 [↑](#footnote-ref-2)
3. <https://www.independent.co.uk/news/business/news/gender-inequality-man-woman-switch-names-week-martin-schneider-nicky-knacks-pay-gap-a7622201.html> [↑](#footnote-ref-3)
4. Bertrand, Marianne, and Sendhil Mullainathan. 2004. "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination." American Economic Review 94 (4): 991–1013. [↑](#footnote-ref-4)
5. Office of the Privacy Commissioner of Canada. (2020, August 13). Pipeda Fair Information Principle 5 – limiting use, disclosure, and retention. https://www.priv.gc.ca/en/privacy-topics/privacy-laws-in-canada/the-personal-information-protection-and-electronic-documents-act-pipeda/p\_principle/principles/p\_use/ [↑](#footnote-ref-5)
6. Québec privacy consent guidelines: Key requirements, impacts and next steps: Insights: Torys LLP (no date) Torys. Available at: https://www.torys.com/en/our-latest-thinking/publications/2023/12/quebec-privacy-consent-guidelines (Accessed: 05 October 2025). [↑](#footnote-ref-6)
7. Institute for Business Ethics ibebriefing58 business ethics and artificial intelligence [↑](#footnote-ref-7)