

**DEPARTMENT: Computer Science**

**Module Code: AI Ethics & Applications**

**Academic Year: 2024**

**Submission Deadline: 07/02/2024**

**Actual Submission Date: 07/02/2024**

|  |
| --- |
| **Student Registration Number** |
| 23030 |

**Declaration**: *I have completed and submitted this work by myself without assistance from or communication with another person either external or fellow student or any AI type of content generator. I understand that not working on my own will be considered grounds for unfair means and will result in a fail mark for this work and might invoke disciplinary actions. This piece of assessment will be continuously checked for its academic integrity until my graduation and the mark will be revised if it is found to breach the unfair means policy. It is at the instructor’s discretion to conduct an oral examination which will result in the award of the final grade for that particular piece of work.*

Abstract:

Using Artificial Intelligence (AI) in HR tasks, like in Strategeion's PARiS system, brings up important moral, social, and legal issues. In this report, all the ethical problems related to PARiS are carefully looked at. The study uses different sources to come up with detailed rules for how Strategeion can use AI in HR practices in an ethical way. At the heart of this approach is the idea that AI systems need to be supervised and changed by humans. Strategeion can help create a future where AI technologies give people more power, make businesses more efficient, and work for the good of society as a whole by promoting a culture of ethical innovation and following the guidelines in this study.

Contents

[1. Introduction: 4](#_Toc158217033)

[1.1 Analysis of moral, social, and legal issues: 4](#_Toc158217034)

[1.2 Moral Concerns, how AI Makes Decisions: Discrimination: 4](#_Toc158217035)

[1.3 Effects on People: 5](#_Toc158217036)

[1.4 Topics of law: Following the rules against discriminating: 5](#_Toc158217037)

[2. A Look at the Development of PARiS: 6](#_Toc158217038)

[2.1 No one controls automation: 6](#_Toc158217039)

[3. Advice on how to avoid similar problems: 7](#_Toc158217040)

[3.1 Being Open and Easy to Understand: 8](#_Toc158217041)

[4. Implementation in the business: 9](#_Toc158217042)

[4.1 Fortifying Ethical Governance: 9](#_Toc158217043)

[4.2 A Way for Users to Give Feedback: 10](#_Toc158217044)

[4.3 Following the law: 10](#_Toc158217045)

[5. Conclusion: 11](#_Toc158217046)

# 1. Introduction:

The introduction sets the scene for the ethics investigation of Strategeion's PARiS system by showing how AI technologies have changed hiring processes and making business decisions.

## 1.1 Analysis of moral, social, and legal issues:

When Hara was turned down a job based on factors that had nothing to do with the job, it made people wonder if AI-driven decisions are fair and include everyone. As machines learn way faster than humans, it needs to be clarified why and how should the use of AI benefit people, companies and businesses to grow. The report talks about what happens when you depend on subjective factors and indicates how important it is to make sure that the criteria used to make decisions are in line with fairness and equality principles. Making decisions with AI: Concerns about openness, responsibility, and the need for human oversight in AI driven processes are raised by the automated rejection of applicants. So, what the report aims to say is that AI without human brain not only doesn’t benefit, but also can harm in some cases what exactly happened in Hara’s case when she applied her job application. ([1])

## 1.2 Moral Concerns, how AI Makes Decisions: Discrimination:

It's necessary to think about moral, social, and legal aspects when using artificial intelligence for HR jobs, like in Strategeion's PARiS system. This study goes over all of PARiS's moral issues, including unfair algorithms, privacy issues, being open, and following the law. Using different sources, the study's goal is to come up with specific rules for how Strategeion can use AI in HR in a good way. The beginning of the study of Strategeion's PARiS system's ethics sets the scene by showing how AI has changed the way how businesses hire people and make decisions. The fact that Hara was rejected for the job because of things that had nothing to do with it, raises questions about how fair and open AI-based decision making is. One should not rely on only subjective factors. The report goes into detailed view about this and shows how necessary it is to make sure that the criteria used to make choices are fair, right and equal. Making a choice with the help of AI make a lot of people concerned about who is responsible, how open things are, and whether AI-driven processes need human failure when applicants are instantly turned down. A balanced approach is shown in the study. Technology does the most of the work while also making sure that moral choices are made and that people can use it when they need to. ([2])

## 1.3 Effects on People:

The fact that Hara was rejected shows how decisions made by AI can affect society and how important it is to take into account each person's situation when hiring. The report shows that companies using AI have a social responsibility such as a reevaluation of ethical priorities and being inclusive. Following the laws against discrimination looks into how well PARiS follows anti-discrimination rules and shows what could happen legally if candidates are turned down because of things that aren't related to the job. This shows how important it is to make sure that hiring methods that are based on AI are in line with the law in order to avoid legal problems and make things fairer. In AI-driven hiring, transparency and responsibility become very important legal issues. People are telling Strategeion to be open and responsible so that applicants can believe them and so that they follow the law and ethical standards. Overall, the study indicates how difficult it is for AI to work with HR practices when it comes to moral, social, and legal issues. Companies like Strategeion need to take into consideration to deal with these problems in a correct way to support fairness and equality, and encourage openness and responsibility in the use of AI. In the end, AI systems must be designed and used with ethics in order to be sure they are on the same page with law and social norms. With the help of automation, it must be ensured that moral decisions are made in a right way and that people can use it in a correct and honest way when need. ([3], [4])

## 1.4 Topics of law: Following the rules against discriminating:

The report looks at how well PARiS follows anti-discrimination laws and shows what could happen legally if candidates are rejected because of things that aren't connected to the job. It shows how important it is to ensure that the hiring procedures that are based on AI are in line with the law in order to avoid legal issues. When AI is used to hire people, transparency and responsibility become very significant legal issues. Strategeion must be open and responsible if it wants to build a company with trusted and right people and follow the law and ethical standards. In the end, the study shows how difficult the moral, social, and legal issues are that come up with AI-driven HR practices. It asks companies like Strategeion to take the initiative to deal with these problems in a responsible way that will support fairness and equality, and encourage honesty and responsibility in the use of AI. It shows how important it is for Strategeion's HR processes to be open and responsible, which are the most important things for creating a culture of trust and obeying the law. ([5])

# 2. A Look at the Development of PARiS:

The creation of PARiS shows a huge defect because of biased training data correlations, mainly the powerful connection between military service and sports involvement. This shows how the model can identify if the candidate is a good fit, which keeps biases alive in the way it makes decisions. The data shows a strong correlation between military service and sports participation, the model by a mistake connects unrelated factors as signs of a candidate's fitness. This kind of systemic bias makes the hiring process less productive and raises ethical worries about what can happen when AI is developed. In order to fix the problem, Strategeion needs to look at its training data as much so it can reduce errors by taking actions. This will help make the AI system fairer and objective, in parallel with good hiring processes. The lack of variety in training data, especially the fact that applicants with disabilities were not considered, makes it easier for PARiS to make unfair decisions. To teach AI systems to understand and take into account a wide range of experiences and qualifications, it must need a lot of different datasets in order to come up with a good result. As the case with Hara, a lack of variety in the training data leads to a limited review of candidates. If the system doesn't have variety of applicants, it won't be able to understand different backgrounds, skills and points of view, which could lead to unfair results. Strategion must make variety of training data a top priority by collecting datasets that show a wide range of experiences, such as people with disabilities. ([6], [7])

## 2.1 No one controls automation:

The fact that PARiS is used without enough human control and too much dependence on automation shows how important it is for AI driven HR decisions to take a balanced approach. AI systems can make things done in a more efficient, faster and better way, but they lack human sense, which makes it harder to analyze and to solve tough problems. Strategions’s failure to predict and address ethical issues during the development of PARiS shows a huge lack in the integration of ethics. Without strong ethical frameworks, the system is open to accidental consequences and unfair results. This shows the importance to include ethical factors when creating AI systems. Strategeion can make sure that AI systems like PARiS follow principles of fairness, transparency, and responsibility in decision-making by combining human mistakes and building in ethical frameworks from the start. Relying too much on automation without enough human oversight can cause unexpected ethical problems. AI systems are fast and efficient, but they don't have human sense, which makes it harder to figure out how to solve complicated problems. This shows how important it is to include ethical considerations from the start.

Strategy can lower these risks by making sure that AI systems like PARiS follow fairness, openness, and accountability rules when making decisions. This means using human knowledge to help automation and building ethical standards into the creation of AI from the start. ([8])

# 3. Advice on how to avoid similar problems:

Strategion's priority must be the usage of a variety of datasets when creating AI models. One of the most important things that can be done to get rid of biases and errors in AI systems is to be sure that people from all backgrounds, experiences, and with different skills are represented. The call for diverse training data comes from the knowledge that the dataset's make-up has a big effect on how well the AI model understands different qualities and qualifications. This effective approach tries to keep biases by showing the AI system a wide range of factors that affect a candidate's suitability for a specified job. Putting diversity first in training statistics is also the right thing to do from an ethical point of view and encourages decisions that include everyone. As a company that values justice and equality, Strategeion can lead the way in making sure that AI is used in hiring in an ethical way by making sure that the models which are trained on datasets that represent the diversity, contain the real life examples. This attention to diverse training data shows that Strategeion is seriously concerned about hiring people in an honest and right way. It also makes the company look like a responsible and welcoming AI user in the hiring process. Therefore, training data should be looked over and updated regularly. It is important to make it a habit to be up to date and review training data regularly, so that AI models can avoid unpredictable errors and don't reinforce biases and stay in line with changing social norms and standards. Society's values change over time, so keeping the moral integrity of AI systems requires being on the lookout. Regular reviews are necessary as they can find and fix biases that may have developed over time without being meant to. As social values change, AI models may become out of date and accidentally strengthen biases that are against ethical standards that we have today, if they are not checked periodically. Patterns or connections that may have formed in the training data can be found through regular reviews. This lets companies like Strategeion to take action and modify their AI systems as needed. Also, training data needs to be updated on a regular basis to keep up with changing social standards. As ideas, beliefs, and standards for inclusion change, AI systems need to be able to adjust and reflect these changes. By adding the newest information to the training data, companies can make sure that their AI models stay up to date on modern ethics issues and continue to make decisions that are fair, unbiased, and in line with what society expects. As part of its dedication to using AI in an ethical way, Strategeion should make it standard practice to review and update training data on a regular basis. This not only protects against unconscious biases, but it also shows that the company is serious about using AI responsibly. By taking this proactive method, Strategeion can show leadership in ethical point of view, setting an example for the responsible and flexible use of technology that fits with changing social values. To mitigate the risks that come with artificial decisions that aren't checked, Strategeion needs to build a strong system for human review into its AI-driven processes. Adding new features that can let people to check and fix possible biases, mistakes, and delicacy in the situation is a part of this. By encouraging human errors, Strategeion can uphold fairness and equity principles. This makes it less likely that unfair results will happen and shows trust in its HR practices. Finding the Right Balance Between Automation and Human Decision-Making In AI-driven HR systems, it's very important to find the best mix between automation and human judgment. Automation can make things more efficient and help with scaling, but people still need to handle complicated cases, especially ones that are unclear, subjective, or involve moral issues. By keeping humans involved in decision-making, Strategeion can make sure that its AI systems work in an ethical and effective way, creating a fair, open, and responsible workplace. ([9], [10], [11])

## 3.1 Being Open and Easy to Understand:

Strategeion should be in an open communication to give clear, understandable accounts of how its AI systems make decisions. Transparent information is a key point of building trust with applicants and making sure they fully understand the factors during the selection process. AI algorithms can be hard to understand, but people need to understand. Clear answers help with this. People who are applying for jobs have a right to know how decisions are made. As part of its dedication to ethical AI, Strategeion can meet this expectation by explaining the main criteria, factors, and things that go into the decisions that its AI systems make. Transparency not only gives applicants what they want, but it also makes the hiring process more honest and responsible. By making the decision making processes less mysterious, Strategeion shows that it cares about being open and fair. Therefore, this gives applicants faith in the AI system's fairness and reduces the risks for possible biases. Clear answers not only show trust, but they also give applicants the power to improve their chances of being chosen by giving them a better understanding of the requirements. If applicants know exactly what qualities and traits the AI system looks for, they can better tailor their applications, which makes the hiring process more fair and open to everyone. Strategeion's promise to give clear answers is in line with the moral imperative of being open about how AI is used. Strategeion is not only promoting trust and fairness by taking this approach, but it is also adding to the larger conversation about using AI in hiring and employment in a responsible and accountable way. Set up clear channels of communication It is very important to set up clear lines of contact so that applicants whose applications are affected by AI decisions can voice their concerns. Having open and honest conversations is key to lowering the worries and doubts that might come up when AI is used in decision-making. In turn, this helps build and keep trust by showing that an organization is willing to answer questions and explain why decisions were made the way they were. ([12])

# 4. Implementation in the business:

All the time it is important for workers who use AI to make decisions to get ongoing training to make sure they stay up to date on possible biases and ethical issues. Because AI technologies are always changing, people who use them need to keep training to keep up with the new social issues and problems that come up. Continuous training is a proactive way to give employees the information and skills they need to find biases in AI systems and fix them. The field of AI ethics is always changing, so employees need to keep up with the latest news so they can spot possible problems and make smart choices. This ongoing schooling is especially important now that AI technologies are improving so quickly and there may be complex moral problems that come up. By learning more about ethical issues on a regular basis, employees are better able to make choices that are in line with company values and societal norms. StrategyOn is committed to ongoing training for workers who make decisions using AI. This makes the company a leader in responsible AI use. By investing in their employees' ongoing professional growth, Strategeion makes sure they have a workforce that is not only good at using AI technologies but also committed to making decisions in the most ethical way possible. Make sure people know how to use AI responsibly. Strategy must actively work to create a company mindset that values responsible AI use above all else. To do this, all workers need to have a common understanding of the ethical rules that govern uses of AI technologies. Creating a culture of responsible AI use means getting everyone in the company to share a commitment to ethical issues. Every person who works at Strategeion should not only know about the social rules that govern the use of AI applications, but also fully support them. This cultural focus on responsibility makes sure that AI technologies are used with a deep understanding of the possible outcomes. This encourages ethical decision-making at all levels of the company. A culture of responsible AI use also makes employees feel like they share duty and accountability. By following ethical rules themselves, team members become active supporters of ethical AI practices, helping to create a place where ethical concerns are naturally part of daily work. Strategeion's devotion to creating such a culture shows that it is serious about being an ethical leader in the use of AI technologies. This approach not only protects against unintended outcomes, but it also makes Strategeion a leader in the AI business when it comes to promoting responsible AI practices. Strategeion can help a larger shift in society toward the responsible and careful use of AI technologies by encouraging a collective attitude that puts ethical concerns first. ([13])

## 4.1 Fortifying Ethical Governance:

As a crucial check and balance system, this group is made up of experts from many fields, such as AI ethics, HR, the law, and the social sciences. Strategeion makes sure that AI applications are fully looked over by bringing together experts with a range of backgrounds and points of view. This is especially important in sensitive areas like HR decision-making. It also acts as a safety measure to make sure that AI uses are in line with the company's morals, ethics, and legal requirements. This multidisciplinary method shows that Strategeion is truly concerned on making a close eye on and being responsible for all of its AI applications and systems. When bringing experts from different fields, the company gets plenty of points of view to be sure that its AI-powered processes are in line with business values, ethical standards, and the law. ([14])

## 4.2 A Way for Users to Give Feedback:

By asking for feedback, Strategeion not only shows that it is committed to being open, but it also gets useful information that can be used to keep getting better. This feedback loop lets the group quickly address applicants' concerns, which creates a place where everyone feels welcome and able to work together. Adding this kind of method not only improves the performance of the AI system, but it also shows that Strategeion is committed to hiring people in an honest and applicant-focused way. The information gathered from applicant’s comments is very helpful for improving and making AI systems work better. By actively including feedback in the development process, Strategeion makes sure that its AI models change along with how applicants' needs and expectations change. This feedback-driven version not only makes AI systems more accurate and useful, but it also shows that Strategeion is still committed to hiring people in an honest and fair way. The addition of a feedback system not only shows that Strategeion is committed to being open and honest, but it also helps them keep getting better and to grow. Including user comments in the development process also shows that Strategeion is committed to hiring people based on their skills and qualifications. Strategeion makes sure that its AI systems change with the correct needs and expectations by giving top priority to applicants' views. This method not only makes AI systems more accurate and useful, but it also shows that Strategeion is still committed to hiring people in an honest and fair way. ([15])

## 4.3 Following the law:

Making sure that anti-discrimination laws and rules are followed in its AI-driven HR practices, Strategeion places a high value on following anti-discrimination laws and rules. Laws say that the company's AI systems must follow them. This makes sure that there is fairness, equality, and nondiscrimination in all parts of hiring, managing people, and making decisions. By taking legal issues into account ahead of time, Strategeion can lower risks, build trust with partners, and create a workplace where everyone has the same chances. Along with making sure that anti-discrimination laws and rules are followed, Strategeion is also very much committed fairness and equality. Moreover, putting legal compliance first shows that Strategeion is committed to creating a workplace where everyone is treated equally and has the same chance. ([16])

# 5. Conclusion:

Looking into Strategeion's PARiS system shows how complicated the ethics issues are when AI is used in HR tasks. We looked at a lot of different sources and used the ethical standards outlined in the IEEE literature to help us figure out the complex things that need to be thought about before AI is used in HR situations. The suggested rules give Strategeion a way to deal with the social problems that come up when AI makes decisions. To make sure that AI-driven processes stay sensitive to the specifics and circumstances of each case, it is very important to find a balance between technology and human judgment. Setting up things like Ethics Review Boards and regular audits also shows that Strategeion is committed to always getting better and following ethics standards. Also, making sure that anti-discrimination laws and rules are followed is a strong reminder of Strategeion's duty to uphold basic ideals of fairness and equality. Finally, getting to ethical AI means always thinking about things, changing things, and working together. By committing to ethical AI practices, Strategeion not only protects against possible harms, but also makes way for innovation that is based on morals and in line with society values. As Strategeion starts this journey, it has the chance to set new standards for how AI can be used in human resources and other areas in an ethical way.

**References:**

[1] Diakopoulos, N. (2016). "Algorithmic Accountability: A Primer." Data Society Research Institute. Available: https://shorturl.at/mqrEO

[2] Brynjolfsson, E., & McAfee, A. (2014). "The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies." W. W. Norton & Company. Available: https://shorturl.at/mxM16

[3] Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). "Semantics derived automatically from language corpora contain human-like biases." Science, 356(6334), 183-186. Available: https://shorturl.at/dfhMQ

[4] Jobin, A., Ienca, M., & Vayena, E. (2019). "The global landscape of AI ethics guidelines." Nature Machine Intelligence, 1(9), 389-399. Available: http://ecocritique.free.fr/jobin2019.pdf

[5] Barocas, S., Hardt, M., & Narayanan, A. (2019). "Fairness and Abstraction in Sociotechnical Systems." arXiv preprint arXiv:1901.10002. Available: https://arxiv.org/pdf/2105.05779.pdf

[6] Chen, I. Y. H., Daumé III, H., & Kira, Z. (2018). "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead." Nature Machine Intelligence, 1(5), 206-215. Available: https://shorturl.at/krCWZ

[7] Zou, J., Schiebinger, L., Hjort, J., & Schaefer, A. W. (2018). "AI can be sexist and racist—it’s time to make it fair." Nature, 559(7714), 324-326. Available: https://shorturl.at/grA56

[8] Buolamwini, J., & Gebru, T. (2018). "Gender shades: Intersectional accuracy disparities in commercial gender classification." Proceedings of the 1st Conference on Fairness, Accountability and Transparency, 77-91. Available: https://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf

[9] Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). "Machine Bias." ProPublica, 23. Available: https://www.science.org/doi/full/10.1126/sciadv.aao5580

[10] O'Neil, C. (2016). "Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy." Crown. Available: https://rb.gy/qxi3jh

[11] Raji, I. D., & Buolamwini, J. (2019). "Actionable auditing: Investigating the impact of publicly naming biased performance results of commercial AI products." Proceedings of the Conference on Fairness, Accountability, and Transparency, 335-344. Available: https://rb.gy/6oa3ok

[12] Diakopoulos, N. (2016). "Algorithmic Accountability: A Primer." Data Society Research Institute. Available: https://rb.gy/43h7ql

[13] Jobin, A., Ienca, M., & Vayena, E. (2019). "The global landscape of AI ethics guidelines." Nature Machine Intelligence, 1(9), 389-399. Available: http://ecocritique.free.fr/jobin2019.pdf

[14] Mittelstadt, B. D., & Floridi, L. (2018). "The Ethics of Big Data: A Map." Journal of Responsible Innovation, 5(2), 186-200. Available: https://rb.gy/zdvqfa

[15] Chouldechova, A. (2017). "Fair prediction with disparate impact: A study of bias in recidivism prediction instruments." Big Data, 5(2), 153-163. Available: https://pubmed.ncbi.nlm.nih.gov/28632438/

[16] Diakopoulos, N. (2016). "Accountability in Algorithmic Decision Making." Communications of the ACM, 59(2), 56-62. Available: https://www.brandonpardekooper.com/automated-decision-making/