Importance of Ethics in field of Artificial Intelligence

Introduction:

An algorithm is a finite sequence of well-defined instructions that describes in sufficiently great detail how to solve a problem. Companies have started using sophisticated algorithms to automate many of their tasks. Algorithms have no ability to deal with grey areas or ambiguity and this is where the risk comes in for any business. The implementation of AI in the task such as hiring process must be carefully monitored by the people behind the technology to avoid the bias. AI is only as smart and accurate as the fed historical data. Using AI in recruitment comes with some unintended consequences. One such case is Amazon recruiting tool.

Automation has been key to Amazon's e-commerce dominance, be it inside warehouses or driving pricing decisions. Amazon team has been building algorithms since 2014 to review the job applicants resume. They trained model to vet applicants by observing patterns in resumes submitted to the company over a 10-year period, rating them from one to five stars. In 2015 the company realized that model was not rating the resumes in a gender-neutral way. Most trained data came from men resume, a reflection of male dominance across the tech industry. so the algorithm learned to systematically downgrade women's CV's for technical jobs such as software developer. It penalized women's resumes that included the word "women's", as in "women's hockey club captain". And it downgraded graduates of two all-women's colleges (Dastin, 2018). Amazon team edited the program to make them neutral to these particular terms but was in vain. so the company has scrapped its recruitment tool at the end of 2017.

The Possibility of creating intelligent recruiting tool raises lots of ethical issues such as Cyber-Snooping, withholding Feedback, Predicting biased outcomes and Black-box selection (Chamorro-Premuzic, 2018). The Liffick analysis can be used to analyze the ethical issue of the above case.

Liffick Analysis:

A. Main Participants and Actions

The amazon gender biased recruiting tool case could have had the following participants.

1. Primary participants:

⇒ Project Manager

- Failed to thoroughly explore the ethical consideration implicit in the design
- Developed the project plan and managed the deliverables .
- Approved team to use the historical data to train the model
- Failed to monitor the bias in the earlier phase of the project

⇒ Human Resource team

- Agreed to provide the data to train the model.
- Failed to notify the project team about the ethical issue such as contextual integrity, fairness involved

⇒ Machine Learning Engineers

- Contributed to the overall project objectives and designed the model.
- Documented the process
- Worked closely with the project manager and provided the expertise.

⇒ Company Management

- The company failed to ensure that their understanding of value-based system innovation is based on de jure and de facto international human rights standards
- Failed to elaborate further about the experiment.
- Setting requirements for filtering out the resumes
- Approved the project team to carry out the experimental automation process.

- Failed to realize the ethical issue such as diversity, irreconcilability and fairness.

⇒ Reuters

- The international news agency who first reported the companies experiment.

2. Secondary participants:

⇒ The people, whose data has been used to train the tool.
⇒ The participants, whose resume has been used in the recruitment tool as part of the experiment
⇒ Various other members as part of this automation process.

3. Implied Participants:

⇒ Programmers

- Implemented the Machine Learning Engineer design.

⇒ Users (Recruiters)

- Who experimented the recruiting tool in the hiring process.

B. Reducing List through simplifying the assumptions:

It is obvious that everyone wanted it to be an engine which rates the resume and provides the top five for hiring. Every participant's action has been associated with the event remotely. The international agency who reported the company's experiment can be exempted from the list. It can be seen that company acted in a good belief and called out to shut down the experiment, as there was no guarantee that the machine would not devise other ways of sorting candidates that could prove discriminatory. But none of the participants can be eliminated other than news agency, who failed to consider the ethical issue of the experiment.

C. Legal Consideration

Code of ethics serves a role in any field that impacts human lives, such as medicine or engineering. Tech organizations like the Institute for electronics and electrical engineers(IEEE) and the Association for Computing Machinery(ACM) also adhere to the code of ethics to keep the technology

beneficial but no concrete ethical framework exists to guide all the researchers involved in AI's Development (Davey, 2017).

In many areas of automation, machine is not only capable of causing serious harm but they assume responsibilities once reserved for humans.
Though there is no direct legal issue in the above experiment, it has more indirect ethical issues. The company stated that they never used the tool to evaluate the candidates, but it is still unclear whether it affected the lives of people who have been experimented by the tool.

D. Possible actions of the participants

⇒ Project Manager could have

- overseen whether such products actually meet ethical criteria, both when experimented, and considering their evolution after deployment and interaction with other products.
- Studied the design processes to identify moments where engineers and researchers can be encouraged to raise
  and resolve questions of ethics

⇒ Machine Learning Engineer could have

- explicitly presented the empirical evidence of compliance and methodology used, such as data used to train the system, algorithms and components used, and results of behavior monitoring

⇒ Company Management could have

- Created roles for senior-level marketers, ethicists, or lawyers who can pragmatically implement an ethically aligned design, both the technology and the social processes to support such value-based system innovation.
- Empowered employees to raise ethical concerns in day-to-day professional practice, not just in extreme emergency circumstances
- Formed Corporate ethical review boards, or comparable mechanisms, to address ethical concerns in relation to their A/IS research
- clarified the relationship between professional ethics and applied A/IS ethics and help designers, engineers, and other company

representatives discern the differences between them and where they complement each other

E. A Possible justification for the participant actions

Ethics is about how we relate to human beings, how we relate to the world, how we even understand what it is to live a human life or what our end goals of life are. AI is raising all of those questions. It's almost impossible to say what AI ethics is about in general because there are so many applications. But one key issue is what happens when AI replaces or supplements human agency, a question which goes to the heart of our understandings of ethics. AI systems will assume responsibility from humans – and for humans –it's important that people understand how these systems might fail. Codes of professional ethics rest on the basic idea that the professionals have an adequate level of control over their models, but the black box problem makes this difficult. The same applies to the above case where the participants have no control over the model they created, as the model update their behavior making it more difficult for the designers to control and understand the decision-making process. The lack of transparency because of black box problem makes it extremely difficult to construct and encode the code of ethics.

F. Key statements

- **"….** started using sophisticated algorithms to automate the recruiting process…"
- "…Using AI in hiring comes with some unintended consequences…"
- **"**….. intelligent recruiting tool raises lots of ethical issues such as …. Predicting biased outcomes and Black-box selection……"
- "Company failed to ensure that their understanding of value-based system innovation….."
- "….he Failed to thoroughly explore the ethical consideration implicit in the design…."
- **"…** did not realize the ethical issue such as diversity, irreconcilability , and fairness…"

G. List of Questions :

There are number of questions raising when analyzing the case

- Is there a need for more constructive and sustained interdisciplinary collaborations to address ethical issues concerning autonomous and intelligent system?
- Are industry lacking value-based ethical culture and practices?
- Is there lack of empowerment to raise ethical concern?
- Is it the right time for the organizations to examine how to flexibly implement value-based design?
- Are institutional ethics committees under-resourced to address the ethics of R&D in the AI fields?
- Are the algorithms behind the intelligence or autonomous system not subject to consistent oversight?

H. Other models related issue :

The designers have lack of control i.e. Black box problem over these kinds of model when using sophisticated algorithms to automate several works. Consider the Northpointe algorithm that US courts used to predict reoffending criminals. The algorithm weighed 100 factors such as prior arrests, family life, drug use, age and sex, and predicted the likelihood that a defendant would commit another crime. Northpointe's developers did not specifically consider race, but when investigative journalists from ProPublica analyzed Northpointe, it found that the algorithm incorrectly labeled black defendants as "high risks" almost twice as often as white defendants (Revez, 2016). Unaware of this bias and eager to improve their criminal justice system, states like Wisconsin, Florida, and New York trusted the algorithm for years to determine sentences. Without understanding the tools they were using, these courts incarcerated defendants based on flawed calculations.

The Northpointe case offers the preview of the potential dangers of deploying AI systems that people don't fully understand. Current machine learning algorithm runs so fast that no one really knows how they make the decisions, not even the developer (Davey, 2017). Systems learn from the environment and update themselves, making it difficult for the designers to control and understand the decision-making process. This lack of transparency-the black box problem -makes it extremely difficult to construct and enforce a code of ethics.

I. Code of ethics:

Though there is no concrete ethical framework exist to guide researchers involved in AI development. The research ethics such as

- Minimizing the risk of harm
- Obtaining informed consent
- Protecting anonymity and confidentiality
- Avoiding deceptive practices
- Providing the right to withdraw
- Agreed to be honest and realistic in stating claims or estimates based on available data
- agreed to seek, accept, and offer honest criticism of technical work, to acknowledge and correct errors, and to credit properly the contribution of others
- agreed to assist colleagues and co-workers in their professional development and to support them in following the code of ethics.

can be used to frame AI's development within the society broader quest of improving human well-being.

J. **Alternative proposal to provide Ethical solutions**:

⇒ **Positive Ethical Possibilities for AI**

Development in artificial intelligence and machine learning provides some of the most interesting breakthroughs in technology. There are many incredibly promising probabilities-including better medical diagnoses, self-driving cars project to significantly lower the rate of serious traffic accidents, improvements in education, fraud detection in financial sectors service industries, solving the social complex problems and many other areas (Forbes Technology Council, 2018)**.** These applications are highly optimistic in improving the lives of humans.

⇒ **Negative Ethical Possibilities for AI**

Artificial Intelligence generates some of the negative ethical possibilities. This is due to many factors, including the nearly ubiquitous influence it will have over many areas of human lives, and immense power it will give to people who own it and its unpredictable nature, among the other things. AI

will have an immense ethical impact because of the potential harms it can unleash (Gates, 2018). And there some ethical problems concerning privacy. The researchers and developers implementing the AI fail to consider the ethical issue it may cause . Some application may not only harm humans but will be a threat to many people.

⇒ **Compromising solution for Ethical Issue by AI**

AI has become a problem solver for the humans but to ensure Artificial intelligence is in the relative position to benefit humanity. Research and design should be underpinned by ethical and legal norms as well as the methods (IEEE Standards Association, n.d.).Value-based design should become the essential focus for the modern organization.

K. Ethical Theory

⇒ Beneficence:

The moral principle of beneficence says that we should not only avoid harm but also (seek to) do good. The point of experimental technology is, however, there is always the possibility of unknown harm. The researchers in Artificial Intelligence should always keep this in mind.

⇒ Avoid creating or reinforcing unfair bias:

AI algorithms and datasets can reflect, reinforce, or reduce unfair biases. Recognize those that distinguishing fair from unfair biases is not always simple, and differs across cultures and societies. One should seek to avoid unjust impacts on people, particularly those related to sensory characteristics such as race, ethnicity, gender, nationality, income, sexual orientation, ability, and political or religious belief.

L. Conclusion:

Automation has always been the trump card for the amazon's dominance in the industry. Artificial intelligence has been used by the industry to carry out the automation. Value-based design methods put human advancement at the core of artificial intelligence. The organization should encourage researchers/developers to employ value-based design methods to create sustainable systems that are thoroughly scrutinized for social cost and

advantage that will also increase economic value of the organizations. To create AI that enhances human well-being and freedom, system design methodologies should also be enriched by putting great emphasis on internationally recognized human rights, as a primary form of human values. There is no concrete ethical frameworks to guide the researchers/developers in the field of AI but commitment to strong ethical guidelines for research will result in value-based design

Bibliography

Association for Computing Machinery, n.d. *Association for Computing Machinery.* [Online]
Available at: https://www.acm.org/code-of-ethics
Chamorro-Premuzic, T., 2018. *FORBES.* [Online]
Available at: https://www.forbes.com/sites/tomaspremuzic/2018/05/27/four-unethical-uses-of-ai-in-recruitment/#12014b4a15f5
Dastin, J., 2018. *reuters.* [Online]
Available at: https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G
[Accessed November 2018].
Davey, T., 2017. *future of life.* [Online]
Available at: https://futureoflife.org/2017/07/31/towards-a-code-of-ethics-in-artificial-intelligence/?cn-reloaded=1
[Accessed November 2018].
DeepMind, n.d. *DeepMind Ethics and Society.* [Online]
Available at: https://deepmind.com/applied/deepmind-ethics-society/
Forbes Technology Council, 2018. *Forbes.* [Online]
Available at: https://www.forbes.com/sites/forbestechcouncil/2018/03/01/14-ways-ai-will-benefit-or-harm-society/#5c133b314ef0
Future of Privacy Forum, n.d. *Future of Privacy Forum.* [Online]
Available at: https://fpf.org/artificial-intelligence-and-machine-learning-ethics-governance-and-compliance-resources/
Gates, D., 2018. *BYU WHEATLEY INSTITUTION.* [Online]
Available at: https://wheatley.byu.edu/the-ethics-of-artificial-intelligence/#_ftn16
Hauser, R., 2017. *TED Talk-Can we protect AI from biases?.* [Online]
Available at: https://www.ted.com/talks/robin_hauser_can_we_protect_ai_from_our_biases
IEEE Standards Association, n.d. *IEEE Standards Association.* [Online]
Available at: https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ead_methodologies_research_v2.pdf
Revez, R., 2016. *INDEPENDENT.* [Online]
Available at: https://www.independent.co.uk/news/world/americas/northpointe-algorithm-

propublica-biased-black-white-defendants-reoffend-a7106276.html
[Accessed October 2018].
Tufekci, Z., 2016. *TED Talk-Machine intelligence makes human morals important.* [Online]
Available at:
https://www.ted.com/talks/zeynep_tufekci_machine_intelligence_makes_human_morals_more_important?referrer=playlist-talks_on_artificial_intelligen