Analysis of Algorithmic Bias in the Service Platform

A Research Paper submitted to the Department of Engineering and Society

Presented to the Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science, School of Engineering

Shubhi Maheshwari
Spring, 2021

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Signature ______________________________________ Date __________
Shubhi Maheshwari

Approved ______________________________________ Date __________
Sharon Ku, Department of Engineering and Society

STS Thesis: **Analysis of Algorithmic Bias in the Service Platform**

Introduction

In the past decade, the world has endured a massive shift into the digital age, where technology manages, produces, and simulates data and actions from small daily tasks to large national decisions. Specifically in the past few years, scientists have found methods to train technology into self-sustaining and independent entities, allowing devices to act interdependently with humans. This ubiquitous nature of technology has stirred up discussions of who and what should dictate the boundaries of technology. As the world continues to become dependent on algorithms and software to dictate daily tasks, this heavy reliance on inanimate beings can come with biased decision making, generalizations, and other consequences. These biases can cause social inequalities, unfair economic and financial advantages, and lead to the formation of a number of unconscious thought ideologies.

Of the vast array of industries utilizing technology, the service platform, which ranges from blue to white collar jobs in e-commerce, healthcare, and finance among others, has dramatically shifted labor and production to AI bots. From robots taking on roles of consultants, customer support specialists, and financial advisors, to algorithms working to deliver financial trading speculations or create artwork, technology drives the service platform and creates much growth opportunity to improve efficiency across the market. One example of a self-aware artificial intelligence technology used in today's service platform is Amazon's Alexa product line, which combines speech recognition with a vast information database to act as a household's virtual personal assistant. Alexa is powered through speech recording, which then is parsed and rendered to create a response (Marr 2018). Machine learning scientists are responsible for creating these possible skills, which can come with bias as data is parsed and reproduced. These Alexa products not only help with Amazon deliveries and groceries, but also help manage calendars, weather updates, and local events for each individual users' interests. Although very useful to some, Alexa's have the potential of obtaining bias through data collection and repetition. As the AI relies on speech and sounds to assist the user, these sounds vary across households. Now, for example, if some household is biased against a specific race in society, Alexa can pick up this quality through their data training algorithms, and cause upheaval in society. This example shows how algorithmic bias can be detrimental, especially in an industry as extensive as the service platform.

The primary focus of this thesis is to analyze where algorithmic bias can occur in the design stage of software development, specifically through an understanding of designer bias. The analysis will be centered around a specific case for investigation, the *artisanal* retail industry in today's current service platform. Using simplistic methods to simulate designer bias, we will attempt to understand how early stages of development play into bias production when working more interactively with the algorithm itself. These techniques should be helpful to a variety of SaaS companies, when working to train designers on how to design their own AIs, and be aware of the potential formation of bias during development.

Literature Review

Algorithmic bias is a crucial factor in understanding how effective these service platform algorithms are. Much research has been done on the impacts of bias and how to mitigate negative outcomes, from design to execution to maintenance of the system. Whether it be the specific "training data-position", or the iterative prediction models, many algorithms accumulate the bias through the data they are trained with, prior to launch (Sun 2020). According to Sun, it can be understood that bias can be found in algorithms before they even obtain new data from the customers. Other ethical studies have attributed algorithms to collect bias through specific design principles. Kraemer explains how algorithms inherently must be value-laden to some extent, decided by the designer, as much as they try to maintain abstract "thought" (Kraemer 1970). Value laden design becomes amplified especially in the digital service platform, where each algorithm is designed by a select few to address the needs of millions at a time. This may be another major source of bias from the grassroots of development.

Going further into the current research on algorithmic bias, it can be seen that there are many factors involved in understanding the importance of the issue, especially because it is very subjective to each individual. Ethical concerns raised by technology have both epistemic and normative concerns, as they are freehand and unsupervised after a given point of time (Mittelstadt 2016). Acting as humanistic beings, AI can be seen as "social actors". With so many dimensions, AI has formed an actor-network with humans, building and growing to a point where it will not only have general intelligence (such as quick computing, storage, and memory management) but will also use machine learning to perform skills as good as or even better than humans (Reed 2018). From government, banking, and security, to driving, cleaning, and shopping, AI's power to manipulate macro and microdata efficiently is creating an unanticipated dependency for humans. Humans constantly have a need for specialized services, and as this dependency grows, AIs will have the power to change conscious decision making as well.

A recent study from the Brookings Institution introduces the idea of "algorithmic hygiene" as a method to find and mitigate the causes of bias that can occur. Using feedback from both business leaders and engineers, the researchers discover that bias can be found in recruitment tools, facial recognition technology, advertisements. The research methodology claims that one way to mitigate these biases, in current practice, can be to scrutinize results for anomalies through real human monitoring (Turner-Lee 2019). This, in relation to the artisanal product market, can be seen as algorithmic product classification can be monitored by humans to ensure there aren't inconsistencies with the data.

One integral case of algorithmic bias that resulted in adverse effects, was the COMPAS algorithm in criminal justice sentencing. A study from 2016 found that the tool had accumulated bias in risk assessment for resentencing by marking a higher number of black individuals as 'high risk' than their white counterparts, specifically based on the lack of scrutinized historical data collection (Park 2019). Similarly, a study had been conducted regarding advertisement placement for STEM careers, exhibiting blatant gender discrimination in play as more men received the ad than women (Lambrecht 2019). These are prime examples of how algorithmic bias is a consistent social issue and explores a sense of urgency for data scrutiny. Within the service platform, previous cases of similar measure can be used to further analyze in which phase of development process should the mitigation techniques be placed.

After delving deeper into the use of AI and the state of existing algorithmic bias, many studies focus on discovering algorithmic bias post-production. They focus on using mitigation techniques, or data scrutiny *after* certain damages have been done. Instead of inspecting the issue in a retrospective manner, this investigation will focus on understanding how bias, specifically designer bias, is formed and what safeguards can be placed earlier in order to mitigate bias in a proactive manner. Users, designers, and consumers all have different ways of defining and perceiving algorithmic bias, and because there are such a complex group of actors in the case, there are mixed positions on what point of focus the service platform should target these AIs.

STS Frameworks

For this research study, the Actor Network Theory Framework and Mediation Theory will be followed. For this research study, the Actor Network Theory (ANT) Framework will be followed. The ANT framework is a concept that places a group of actors and actants into an interdependent network, where the behavior of the actors correlates to the effects of the actants. This theory can be applied to this thesis investigation, as the designers of the product, shoppers, artisans and small businesses are the actors and the AIs are the actants. When the actors' behavior directly determines what the AI actants learn and model, we can easily see how algorithmic bias can be formed, and how the network explains the robustness of technology adoption in everyday services.

Specifically, in the artisanal service industry case investigation, we can see how the designers, as the actors in ANT, can have biases when building their AI product curators and hence skew the accuracy and ethical efficiency of the technology. To illustrate this network, the service platform is the "network", in which the human entities (actors) recruited are the retailer, the creator, and the designer, and the non-human entity (actant) is the AI. As the human entities provide information, such as retail preferences, price scales, material qualities and other factors, to the non-human actant, the AI can develop skills to mimic the human and speed up the product discovery process. By decreasing search and filtering time by tenfold, product sales will be more specific and thematic. The construction of this heterogeneous network will allow the technology to be designed by societal standards. If designed well, the AI can function as an honest broker and tend to the needs of the stakeholders without compromising accuracy.

As a direct side effect, when the human and non-human entities are connected, bias can be introduced. Humans are inherently diverse-minded and have unconscious biases. For a realistic example, when a mother raises a child, she will have a set of views, principles, and behaviors that she will instill in the child. These characteristics, good or bad, will lead the child to act in a similar way and have similar views and biases. This example demonstrates the effect of the human actors on the AI actants in the service platform network, as training these algorithms to act based on personal views and experiences will intrinsically lead to biases. This artisanal product service platform also has a number of steps from inventory management to payment collection, which will all touch the human to non-human relationship and bias that translates through.

When analyzing the different ethical issues that come with algorithmic bias in the service platform, there are many ethical factors to consider especially in the world of artificial intelligence. Engineers of such a product must take into consideration what consequences could come with the decisions that the entity makes, as a self-sufficient, intelligent being makes decisions based on behavioral data of real people. Last semester, I applied the ANT framework to my case study that created a service network between humans and the AIs in forming a mutually beneficial relationship in retail. This semester, Spring 2021, there are similar approaches we apply with the Mediation theory framework to the research.

The technology mediation framework details the way that technology bridges certain gaps between humans and their reality. It emphasizes that technology plays a major role in creating the interaction between humans and their reality, rather than just acting as an individual entity in the human world. These concepts can be applied to the AIs in artisanal retail and controlling and containing the bias that comes with AIs, as we should not be creating these AIs to simplify the product and make it more scalable. Instead, these AIs should be designed to create simpler interactive realities for humans when shopping, creating and selling products. AIs are specifically designed to indicate human subjects and a human-world relation of virtual shopping and preferences. The framework also mentions that the designers play a major role in designing said "mediation" that facilitates people's perception of reality, moral decisions and practices. In a much narrower sense, if these AIs are designed in a way that begins to lead or direct humans to make certain shopping decisions, have different morals and views on certain products, and do not find the virtual shopping experience fit to their reality, there could be larger effects in play. In this way, it is implied that there is major responsibility for the designers to ensure they address these details and anticipate the behavioral effects caused by the product. So, if this is the case, understanding designer bias when it comes to making an AI or algorithm for the service platform is very important.

Research Method

To conduct the research, three main data components are required for the analysis. First, surveys of the main target audiences, including shoppers, artisans, and small retailers, will be used to understand what kind of types of factors would be considered when making an algorithm for this industry as well as simplistic designer bias from survey formulation. The goal was to gather over 200 shopper, 100 artisan, and five retailer survey responses. The surveyed individuals came from diverse, complex backgrounds, ranging from ages 15-60+, in different socio-economic and racial demographics.

Second, interviews of the computer science designers will be used to understand where bias can or cannot be formed. This will be done through simple image impressions to simulate image recognition bias and hypothetical scenarios to place the designers in the role of the AI. Lastly, the above literature review, particularly on topics such as designer bias in survey analysis, designer bias in product design, designer bias mitigation techniques, and designer bias effects on algorithm design, will be used to consolidate all the findings together, and create a holistic overview on the topic.

To begin the research, interview scripts were designed, and a group of computer science designers (primarily from my Capstone team) were gathered as subjects. Afterwards, each of responses were collected to do a deep dive into what their answers imply, compare their answers against one another, and finally draw some inferential conclusions about where bias can be formed and potentially how to avoid it. In parallel, a question analysis of the surveys will also be completed to gather some concrete findings from that component. This literature review will add on to the initial literature review from STS 4500 and help combine all the findings into one final conclusive report on the topic.

Data Collection

As mentioned above, two main components were used for data collection: surveys and designer interviews. Survey methodology was chosen to imitate the most direct form of designer bias as a product idea and its implications come directly from the designer, and so these surveys helped combine user views with designer views. Interviews were used as a form of personalized 1-1 questioning, seeing how an individual would answer without other motivating factors.

Surveys

Surveys are being used to understand a simplistic form of designer bias that can be formed when formulating questions and design principles. We created a list of questions for each respective target audience (the artisan, the shopper, and the retailer) to gauge customer views on the potential product idea. These responses are not necessarily directly related to the STS thesis topic, however I will be analyzing how these questions were formed and how they could have skewed the overall trends that were drawn from the responses. The surveys were created using Google forms and the survey links are provided in the appendix. Each survey was approximately 12 questions long, to keep it concise and quick to complete. Demographic information collected in the surveys included age, race, and state/country of residence. Other information collected included shop links and product areas.

Interviews

Interviews were conducted on computer science designers/programmers to see where bias can be formed before working with the actual algorithm and its self intelligence training going forward. Again, this is a very high level view of understanding what kinds of things can create bias and the questions will be about the specific artisanal retail industry and service platform. Questions used in the interview can be found in Appendix E. General information about the designers such as their academic background, professional experience, and interests was also collected. Then, certain image classification questions and hypothetical situations were posed to the subjects, similar to those that the product would face when operating through artificial intelligence. The image classification was intended to show how the designers viewed certain products prior to development and what factors they would use to categorize them. The hypothetical situations were used to show certain conditional checks that the AI would need to account for when curating and populating artisanal products to the retailers.

Analysis

After collecting and organizing data through surveys and interviews, it can be seen that a lot of how an algorithm operates, trains, and models data is a direct result from the way that it was constructed. Having certain views on topics and products, along with confounding variables, can create bias within the algorithm unintentionally.

In this case study, surveys were used to understand if there was the possibility of bias in the questions that were presented to the three focus groups; the artisans, the shoppers, and the small retailers. The surveys were gathered from roughly 350 shoppers, 150 artisans, and 2 small businesses, meeting close to the defined goal. The surveys were written by the designers themselves, and basic trends were taken from each of the survey data pools. Trends in the survey were primarily positive, with some negative but changeable key points, and most of the surveyed individuals found that the product platform had potential in the market, giving some concept validation to the team. There were, however, many trends that were drawn from a limited range of answer choices, and questions. For example, when shoppers were answering questions such as "On a scale of 1-10, would you be inclined to purchase an item you couldn't find anywhere else?", the question directs the user to answer on how they shop based on just the factor of "uniqueness and exclusivity in the marketplace." This is not necessarily a negative thing, but rather the question should have been posed to ask what kind of factors are considered when shopping *holistically*. This would be a lot more vague and would not bias the shopper to consider uniqueness at all. Questions like this show that there is bias already built into the survey by the designers. We created questions that would help us directly validate our product by presenting questions specifically asking for trends relating to "unique shopping and products in the artisanal marketplace." This not only skews the results we would have gotten without such specific question phrasing, but shows how designers are already biased when it comes to building a product and surveying for concept validation should be more holistic in nature.

The second method of data collection that was used to understand the idea of designer bias in algorithms was interviewing two computer science designers on my team that would be building the algorithm for the product curation tool. The two interviews had quite different responses and found that a lot of the image classification was interest based and differed between the two engineers. When shown an image of a small plant holder product, subject one answered, "I would not classify this as a popular product, in my opinion, but would put it in the home decor category," while subject two relayed "This product is cool, I would add it to my bookshelf at home." This shows that although the subjects were shown the same image of the same product, both had varying views on the product and would ultimately classify it or priority queue in different ways.

Similarly, the hypothetical situations also showed that there were many external factors that played into how an algorithm should act when showing products to the retailers, along with how the engineers believed each of the target audiences should be treated on the platform. When asked, "How would you respond to a search result showing a certain group of products more than others, and if yes, does this hurt the success of other products on the platform?", subject one answered "I would think that the algorithm has been following my movements on other search pages and would think that it believes that I would be interested in this type of product more. And yes, it could because then we aren't able to show new products to customers and broaden their horizons to all kinds of curated artisan products." Subject two had a slightly different response stating, "Definitely given how we see things on Google, Amazon shopping, and Instagram ads, I would think its data sharing across companies targeting me to see certain products more than others. But, Yes, but it could also make for more successful sales as targeting can be good for making the shopping experience more interesting and direct." Both had different views on what should be prioritized; the customer or the business bringing in profits. These views would definitely play into which groups are emphasized in the actor network, and how the algorithm would be built to prioritize certain groups. The interview responses show that preexisting designer bias can easily be found when building an AI that needs to conduct the same manual human tasks.

To analyze the results, the mediation theory can provide a solid framework to describe the effects of the study. In this theory, the artisanal service platform is a value-laden device in which many factors insinuate biases among shoppers, artisans and retailers. As the technology shifts the way the users behave online and in the marketplace, humans would interpret reality in a different way by modifying their shopping behavior and needs. As bias is created through the technology's preferences and data modeling, and depending on who is benefitting from the technology through initial design, all actors have possible losses. Artisans can be affected by loss of sales and credibility, retailers can lose attractiveness and appeal in the market, and shoppers can lose interest based shopping and personality. The AI technology in place would mediate and prioritize certain users, and would change the way that people shop, by influencing sales of certain products, certain producers, in certain times.

This is all driven by the initial stages of designer perspectives that would determine the backbone and infrastructure of these AI devices and lead it to act and bridge the gap in the artisanal service platform accordingly. For example, if in the case that designers prioritize the retailers when producing the products, artisans would be disadvantaged and forced into additional competition in the retail market. Not all artisans have the same levels of disposable incomes, and so small owners cannot pay the same fees as other producers, creating bias against those who can and cannot afford the services in the market. The overall outcome of the bias creates disparities between sellers and buyers, and disproportionally benefits certain entities.

Conclusion

After conducting multiple data collection experiments, and analyzing trends that followed, it can be seen that much of an AIs algorithmic bias comes directly from the way it is designed. When an algorithm is first developed, for a specific function, tool or customer, there is a vast amount of factors that go into how the algorithm should operate, what functionality it should have, and how versatile it should be. Designers have the ability to dictate what pieces go into the algorithm's functionality and performance, and are also responsible for the data it collects and uses to become a self-sufficient program. This thesis evaluated types of designer bias that can occur prior to the actual development of an AI, using the artisanal retail industry in the service platform as the network in ANT. Using survey data collected from artisans, small retailers and shoppers, the team was able to correlate response trends to question types and wordings. It was found that many of these questions were written in a way that directed the responders to answer in a certain way, rather than have bias-free responses. Although this was quite high level, it is a good example of where and how designer bias can be formulated and baked into the actual design of software. That bias then would be continuously used and recycled when the AI learns to self operate and trains from existing data. Second, the interviews conducted on the computer science designers showed similar results. Each of the subjects responded to artisanal product images and hypothetical scenarios in ways that fit their interests, aesthetic and views. This helped to reinforce the idea that designers must be aware of these aspects when developing AIs for their products, and must have methods in place to safeguard against this.

Combining the literature review research with the high-level findings from the pre-development experiments, we can see that bias will continue to exist, however can be mitigated through efforts in conscious design decisions pre-development. Although there are numerous stages for the AI for bias to be formed post-development, especially if data is not scrutinized often enough, pre-development designer bias is something much more concrete and substantial. The mediation theory shows that customer prioritization must be considered with bias recognition in the market to create a balance. Methods such as using numerous design perspectives, working with diverse minded designers to have varied views, using design psychology principles to remove bias at each new stage of development are all mitigation techniques that can be used to prevent AIs from forming new biases or having preexisting biases as they begin their functions. Future work could include doing an interactive experiment of building the product curation tool mentioned through the investigation, and log certain design decisions, with reevaluation along the way.

Appendix

Appendix A: Surveys

Creator/Artisan Survey → https://forms.gle/hWLLCi2W1xXjEnbP7
Shopper Survey → https://forms.gle/EciUtrAxy21H7JdS7
Small Retailer Survey → https://forms.gle/gAhbbauWN8hdLc1aA

Appendix B: Survey Responses

1. Survey Responses (Artisans + Shoppers)
2. Survey Responses (Businesses)

Appendix C: Survey Trends

https://docs.google.com/document/d/1QSPxva_co8CP1d0A0kQD6tpdHjmGq9EA7Fb1VvOlkSk/edit

Appendix D: Summarized Interviews

Summarized Interview Data

Appendix E: Interview Questions

Interview Questions

1. Ask major, year, experience etc for background demographics
2. Run through a series of images of products that could potentially be sold on the platform and ask for views on them
    1. How would you classify the image?
    2. What would bring this image up in the search feed?
    3. What factors should be associated with this product?
3. Run through a series of hypothetical scenarios that could occur on the platform during use and see how they respond. Examples include - How would you respond to a search result showing a certain group of products more than others?

4. What kind of conditions should be in place to make the market equal opportunistic for the artisans?
5. How does the algorithm play into ensuring the retailer? Etc.

Appendix F: Data Collection Sheet

Data Collection Sheet

Data Source	Stakeholders	Qualitative & Quantitative Data	Themes/Notes	Research Question
Surveys	- Businesses (small retailers) -Artisans -Shoppers	Data was collected from 2 retailers, 150+ artisans, and 300+ shoppers See Appendix B for distinct data that was found.	- able to get a lot more shoppers and artisans to fill out the survey than businesses. See Appendix C for specific list of survey trends. - found most surveyed people answered in similar ways, were limited in choices (scales ranging 1-5) - artisans were biased towards being Esty creators rather than all platforms combined; confused the group in answering	1. How can bias be created prior to the actual production of an algorithm, when trying to define a product? 2. When survey questions are asked looking for insight into how to design the product/algorithm, can these responses be influenced by the designer questions? 3. How much does this impact the actual design of the algorithm?
Interview	CS Designers (2 collected)	Collected 2 1-1 interviews (using question list above) -	Appendix D for Summarized Interview Data	1. Looking to understand how the developers react to certain questions in this product area and how these factors could affect the designer to bake bias into the algorithm

Bibliography

- Glasheen, J. (2020, April 23). The Future of Retail Lies with Artisans, but the Big Guys Can Still Cash In. Retrieved December 03, 2020, from https://www.therobinreport.com/the-future-of-retail-lies-with-artisans-but-the-big-guys-can-still-cash-in/
- Kraemer, F., Peterson, M., Resnik, D., & Turilli, M. (1970, January 01). Is there an ethics of algorithms? Retrieved October 30, 2020, from https://link.springer.com/article/10.1007/s10676-010-9233-7
- Lambrecht, A. (2019, April 10). Algorithmic Bias? An Empirical Study of Apparent Gender-Based Discrimination in the Display of STEM Career Ads. Retrieved October 30, 2020, from https://pubsonline.informs.org/doi/abs/10.1287/mnsc.2018.3093
- Miller and Kartik Hosanagar, A. P. (2019, November 11). How Targeted Ads and Dynamic Pricing Can Perpetuate Bias. Retrieved October 30, 2020, from https://hbr.org/2019/11/how-targeted-ads-and-dynamic-pricing-can-perpetuate-bias
- Mittelstadt, Daniel Brent. (2016)The ethics of algorithms: Mapping the debate - Brent Daniel Mittelstadt, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter, Luciano Floridi, 2016
- Narumitsu, N. (n.d.). ICT and the Future of the Retail Industry - Consumer-Centric Retailing. Retrieved from https://in.nec.com/en_IN/about/technical-journal/latest-isse/year/g15/pdf/150107.pdf
- Park, A., About the Author 2019 UCLA School of Law JD Candidate, Chapter, U., Tsosie, R., Wright, C., & Stevenson, J. (2019, September 21). Injustice Ex Machina:

Predictive Algorithms in Criminal Sentencing. Retrieved October 30, 2020, from https://www.uclalawreview.org/injustice-ex-machina-predictive-algorithms-in-criminal-sentencing/

- Quinn, J. B. (n.d.). The Impacts of Technology in the Services Sector. Retrieved October 30, 2020, from https://www.nap.edu/read/1671/chapter/7
- Reed, M. (2018). The Classification of Artificial Intelligence as "Social Actors". Retrieved from https://scholarworks.gsu.edu/cgi/viewcontent.cgi?article=1059&context=rs_theses
- Sun, W., Nasraoui, O., & Shafto, P. (n.d.). Evolution and impact of bias in human and machine learning algorithm interaction. Retrieved October 30, 2020, from https://journals.plos.org/plosone/article?id=10.1371%2Fjournal.pone.0235502
- Turner-Lee, N., Resnick, P., & Barton, G. (2019, October 25). Algorithmic bias detection and mitigation: Best practices and policies to reduce consumer harms. Retrieved October 30, 2020, from https://www.brookings.edu/research/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer
- Marr, B. (2018, October 5). Machine Learning In Practice: How Does Amazon's Alexa Really Work? Forbes. https://www.forbes.com/sites/bernardmarr/2018/10/05/how-does-amazons-alexa-really-work/?sh=2433f66319