**Analysis of Algorithmic Bias in the Service Platform Prospectus**

A Partial 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 Prospectus

## Introduction

In the service platform today, one of the most prominent market sectors is the artisanal product market, generating over $34 billion in annual revenue globally (Glasheen 2019). Although there is substantial growth potential, the traditional artistic market has historically grown through personal connections, trust-based relationships, and local wholesaling. The artisan network is close-knit, caters to a very distinct audience, and is generally more expensive than mass-produced retail. Similarly, small businesses that sell these goods operate in very traditional atmospheres, with small regional reach and limited marketing capabilities. In addition, in recent years, shoppers have also taken a new interest in the handcrafted market, looking to fill the voids that big retail corporations cannot.

To consolidate all of these limitations, this past summer, my start-up team and I worked on building a software service company that facilitated partnerships between artisans and small businesses in retail to make selling and finding artisan-made goods simpler. To do this, we designed a business model that would leverage technology to carefully curate and scale product discovery in the marketplace to create mutually beneficial relationships between the two target audiences. This social entrepreneurial venture will use technological models and strategic geographical lines to streamline access to products, with the ultimate hope of bringing sustainable, diverse shopping experiences back to society. Starting with simple algorithms, as we continue to grow, we plan on adopting artificial intelligence and vision learning technologies to improve the accuracy of product allocation. However, with new learning models, comes potential bias. The motivation behind this particular study is to analyze and understand how algorithmic bias in the service platform can cause inaccuracy and limitations

of sales, and how this could ultimately affect our creators and their success within the platform.

In recent times, AI technology has dramatically been adopted to simplify complex issues and remove human intervention in the service platform, through the automation and pattern recognition of historical data. As the service platform goes through major shifts, from brick and mortar shopping to the online marketplace, technology inherently faces the issue of accumulating certain biases over time. Algorithmic bias in service platforms is prevalent all over the world. From bias in hiring tools, to messaging platforms, AI learning patterns have been designed to inherently adopt the human biases that were meant to be removed in the first place.

Algorithmic bias in the service platform can be analyzed through two major avenues - the paradigm shift of retail through technology, and the effects of algorithmic bias in this industry. Given the start of the 20th century, the age of the Internet has dramatically changed how people spend and save, with access to e-commerce sales and virtual money management. The retail industry has shifted into a new digital age, where suppliers have shifted power almost entirely to the demand side, and now, must be able to keep up with all virtual scales and fast-paced changes in market trends (Narumitsu 2015). Both of these shifts give way to establishing how technology, now in the AI space, can further increase the productivity of services and how suppliers will have more power to cater to the consumer-centric models if data is manipulated properly. On the other hand, there are still some concerns regarding how data manipulation in the service platform can have negative effects. Although data manipulation targeting is more strategic and sophisticated, a recent study regarding dynamic pricing shows that customers feel they are treated unfairly because the algorithms offer

unequal prices to consumers in different zip codes (Miller 2019). This is a prime example of how algorithms can interpret data in ways not ethical to the common user, and retail is a massive space to incur that controversy. This raises the fundamental questions about how we can maintain ethical, economic, and legal controls when using AI as mediators in the service platform. The primary question is to analyze how algorithmic bias occurs in design and data processing in software service platforms, specifically through an investigation of the artistic goods market.

**Research Question**

As the designers of the technology, we need to investigate where and how algorithmic bias occurs, the ethics and limitations of these design decisions, as well as the impacts on the stakeholders and overall user experience. What would happen if the algorithm develops preconceived notions about the quality of new products, or the individuals selling the products? Could the algorithm develop biases against certain retailers if they have fewer reviews, sales, or product demand? How often should existing data be scrutinized? Visual data, ratings, personal information, and other factors could potentially create certain behavioral tendencies for the technology to follow.

The primary research question is to analyze where algorithmic bias occurs in the design stage of software development, specifically through designer bias.

**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 various stages of the process including 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, pre-launch (Sun 2020). 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). 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 they 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.

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 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 on 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 Framework and Research Method**

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 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 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.

However, as a 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 behavior 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.

After an extensive literature review of sources from SAGE journals, UCLA Law Reviews, and other established journals, and problem identification through the startup's design decisions, I will be conducting two research methods - survey design analysis and interviews and bias simulations. First, I will be surveying a number of individuals in four primary groups - the creators looking to sell their work, the retailers that act as suppliers and would be interested in using the service, the end customers receiving the products and their views on the matter, and lastly the designers of the technology itself. After gathering responses from each of the groups, I will analyze the data collected and determine trends in this market. Afterwards, I will assess the wording of the questions and survey distribution style. This will help determine whether there was designer bias baked into the questionnaires and or if the questions led interviewees to answer in biased ways. Connecting the dots between the survey style and distribution to the trends that were found, we will be able to recognize an early stage of designer bias. The sample size will differ depending on accessibility to each group.

Second, I will conduct psychological bias interviews and ethical simulations on computer science designers. Through simple questions, I will determine which kinds of bias the designers hold for the artistic service platform currently, and what type of conditional checks could be incorporated into the algorithmic design and development to mitigate the potential biases. The number of participants in this study will be limited and used to extrapolate data to understand the technological design effects. These designer interviews will provide a prime explanation of the ANT framework, with the designer (actor) creating an AI product (actant) primarily for one customer, the retailer (actor) can have unanticipated biases against other actors in the network.

## Timeline

The data will be collected over the course of a few weeks. Weeks 1-3 will be used to conduct designer interviews, ethics simulations, and surveys. The following weeks 4-5 will be used to analyze the statistics, draw insights, and construct themes and trends from the data. The final weeks will focus on writing the report and consolidating all information. Milestones will be determined based on numbers of data points collected.

## Conclusion

The research framework will use real life data to understand and analyze the effects of algorithmic bias through the grassroots level of design, and implications across various affected groups in the artistic market of retail. This will be done through an investigation case study involving the artisanal products retail industry, and four experimental groups. Through the analysis of how, where, and why algorithmic bias is formed and found in applications in the service platform, the research will ultimately help illustrate the concept of algorithmic bias in the wide spectrum of SaaS companies, and identify ways to overcome these biases through acknowledgment and practice in design approaches.

## 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