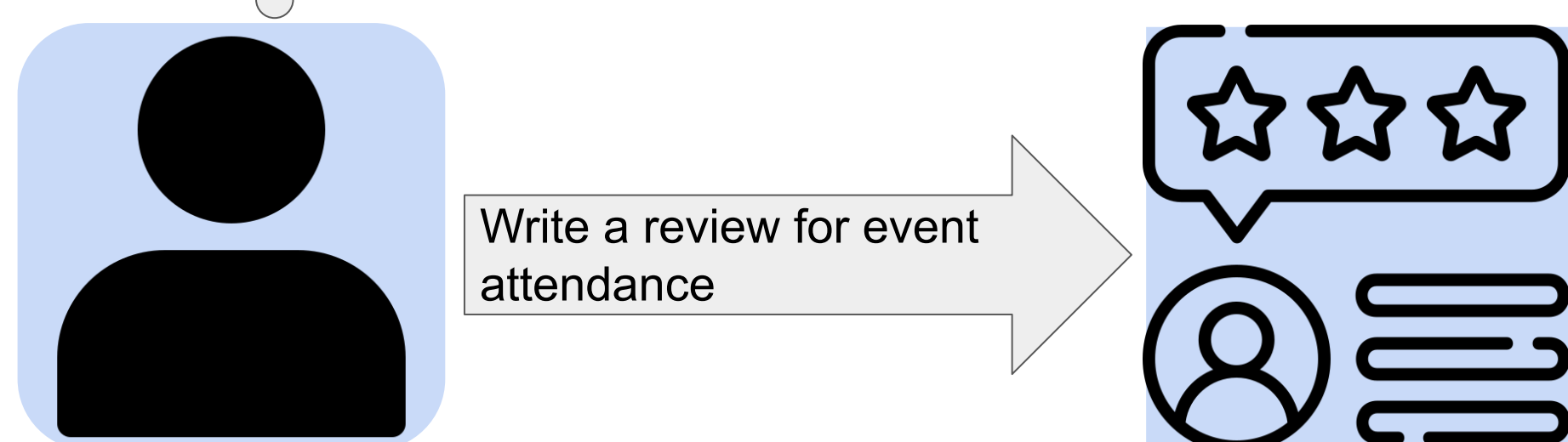
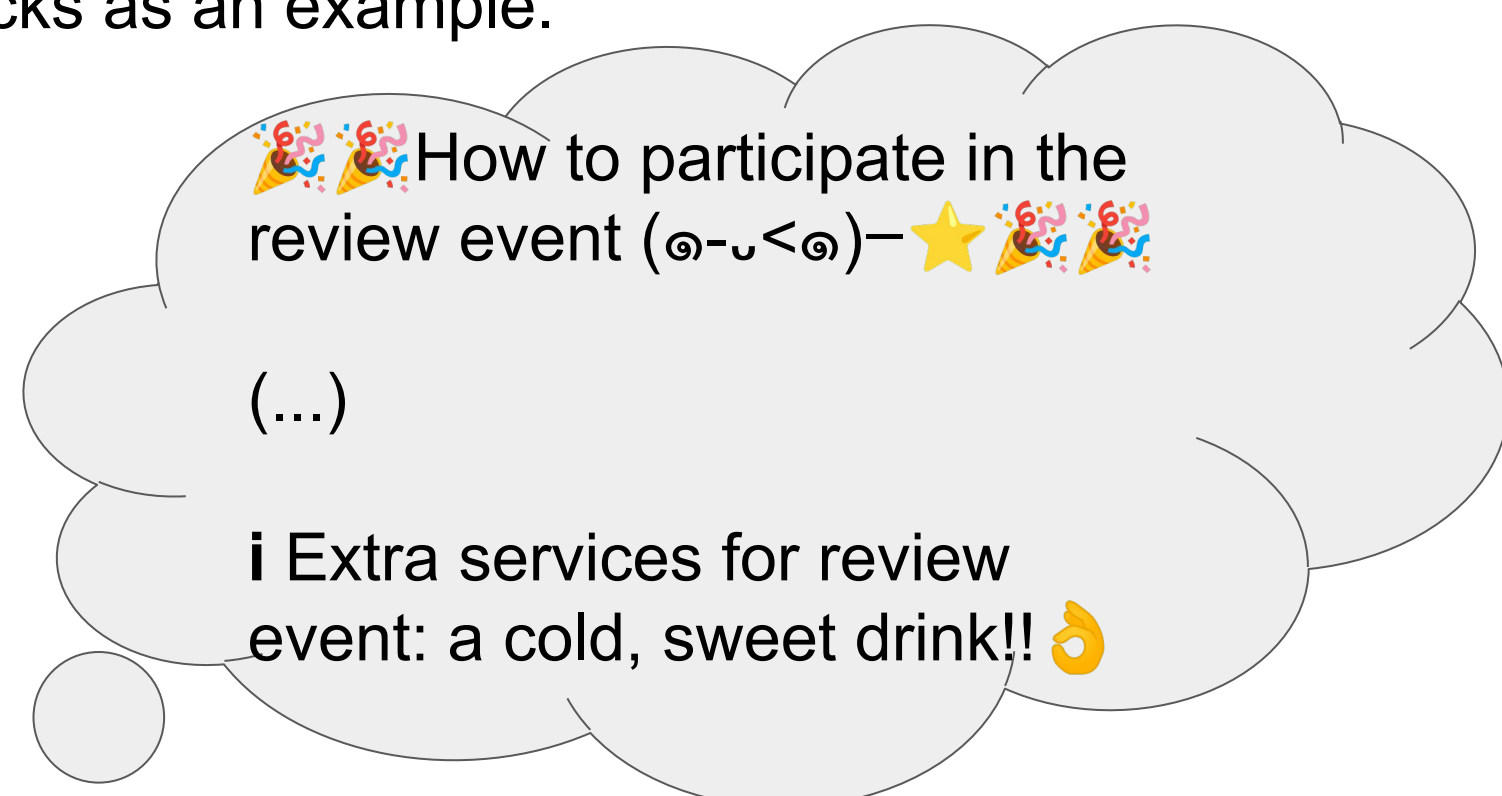
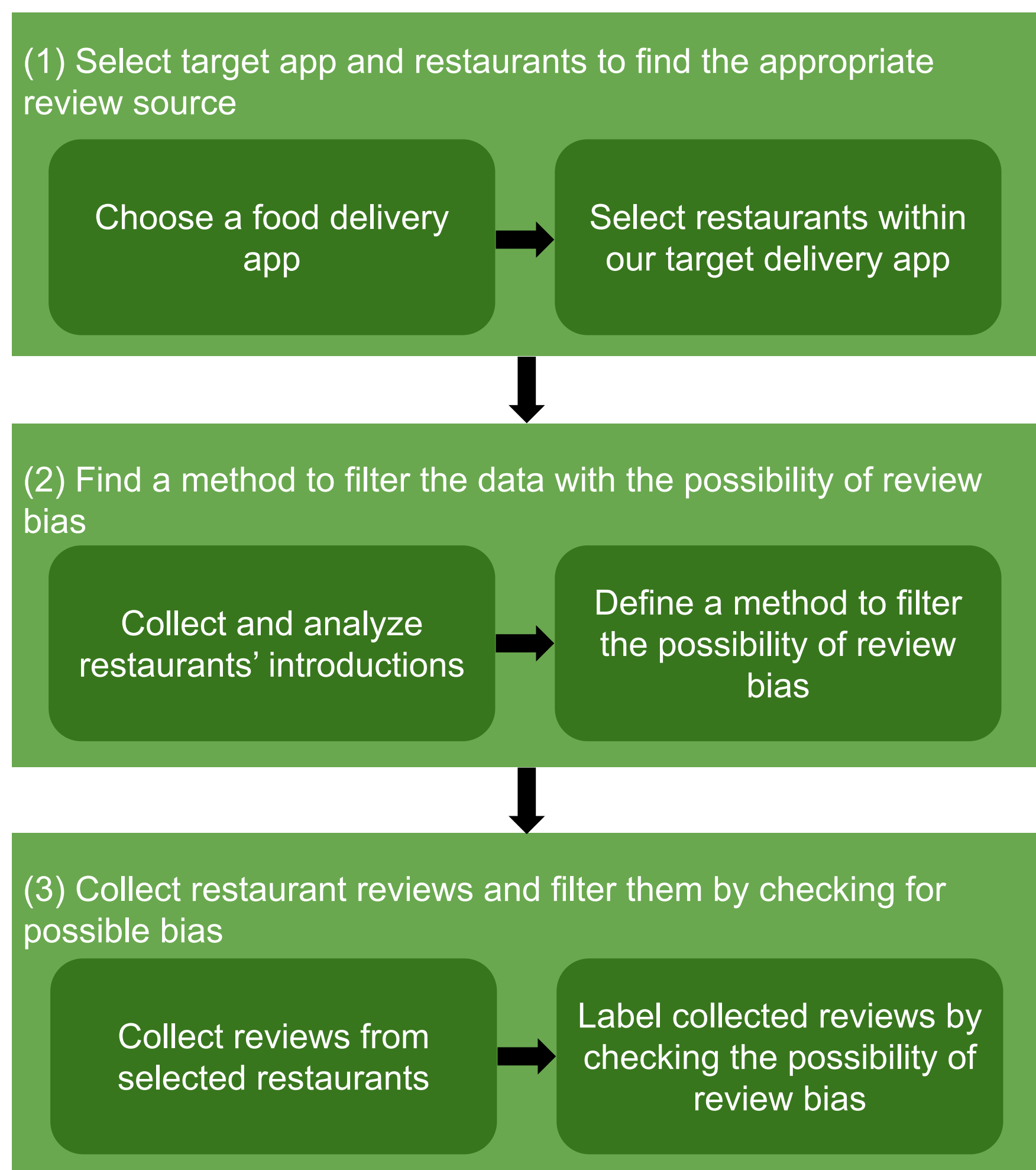


## INTRODUCTION

- The purpose of this study is to present a case of creating a dataset without poisoned data by collecting and filtering reviews in food delivery apps.
- Sometimes, the training data of AI models are intentionally attacked by poisoning, which poisons the training data to drive the learning of AI models in the wrong direction.
- However, previous studies reporting datasets for model training purposes have not considered the possibility of data poisoning attacks for each data.
- Therefore, our goal is to contribute to the community of AI model developers who aim to create fair AI models by proposing a bias-free dataset of food delivery app reviews without data poisoning attacks as an example.

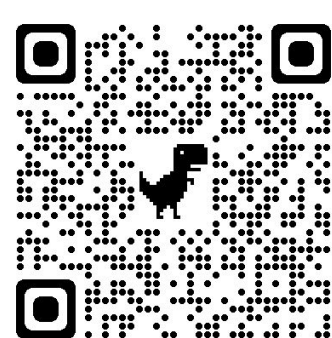


## METHODS



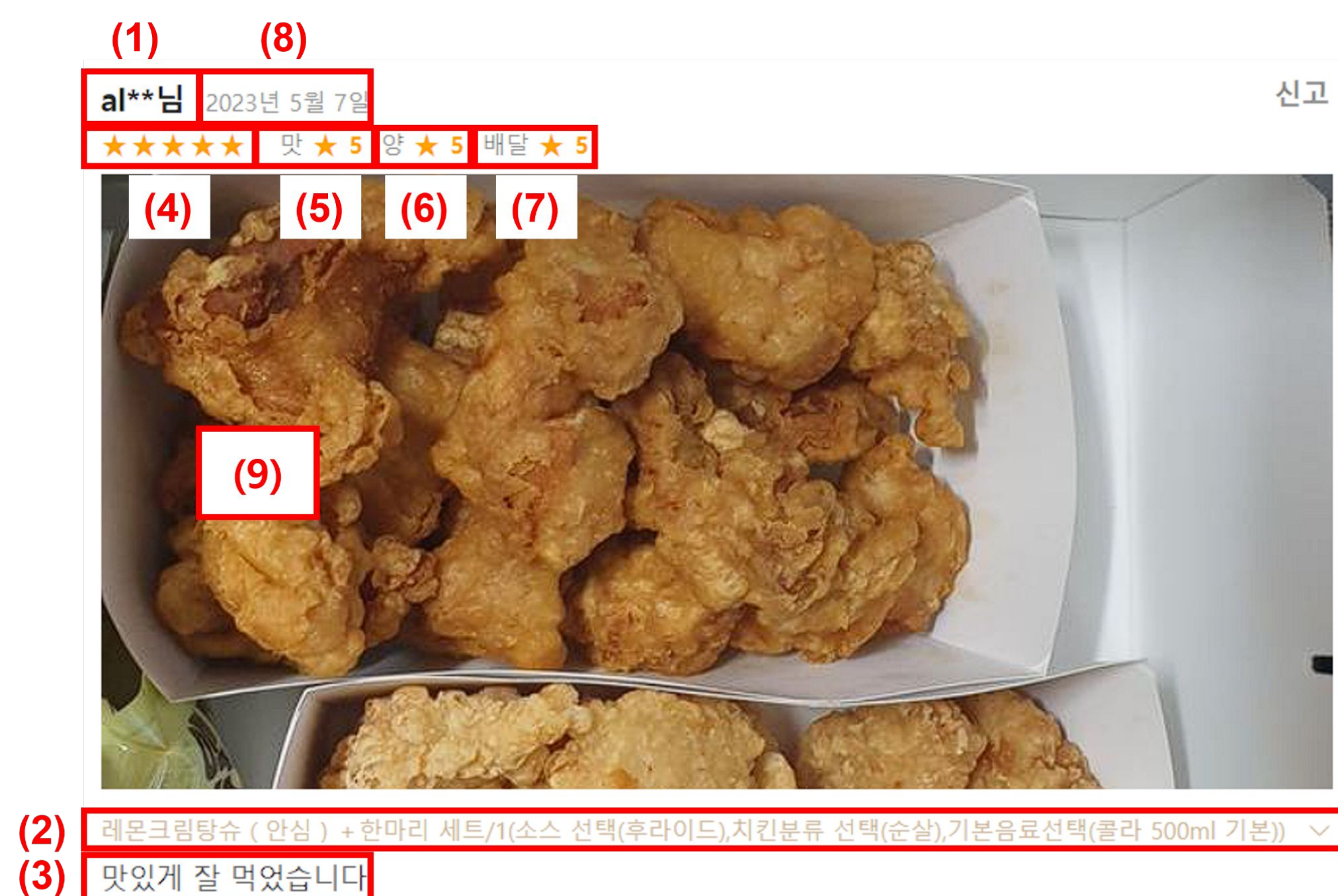
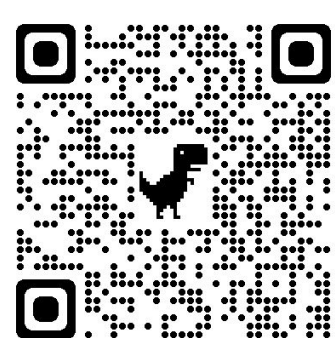
**Figure 1.** The flowchart of our dataset creation process.

- Our goal is to create a dataset without poisoned data.
- First, we selected the target app and restaurants to find the appropriate source of reviews (Figure 1-(1)). Then, we collected reviews from the target restaurants. See Figure 2 for an example.
- Second, we found a method to filter the data with the possibility of bias (Figure 1-(2)). Figure 3 gives more information about our filtering method.
- Third, we collected restaurant reviews and filtered them by checking the possibility of review bias (Figure 1-(3)).

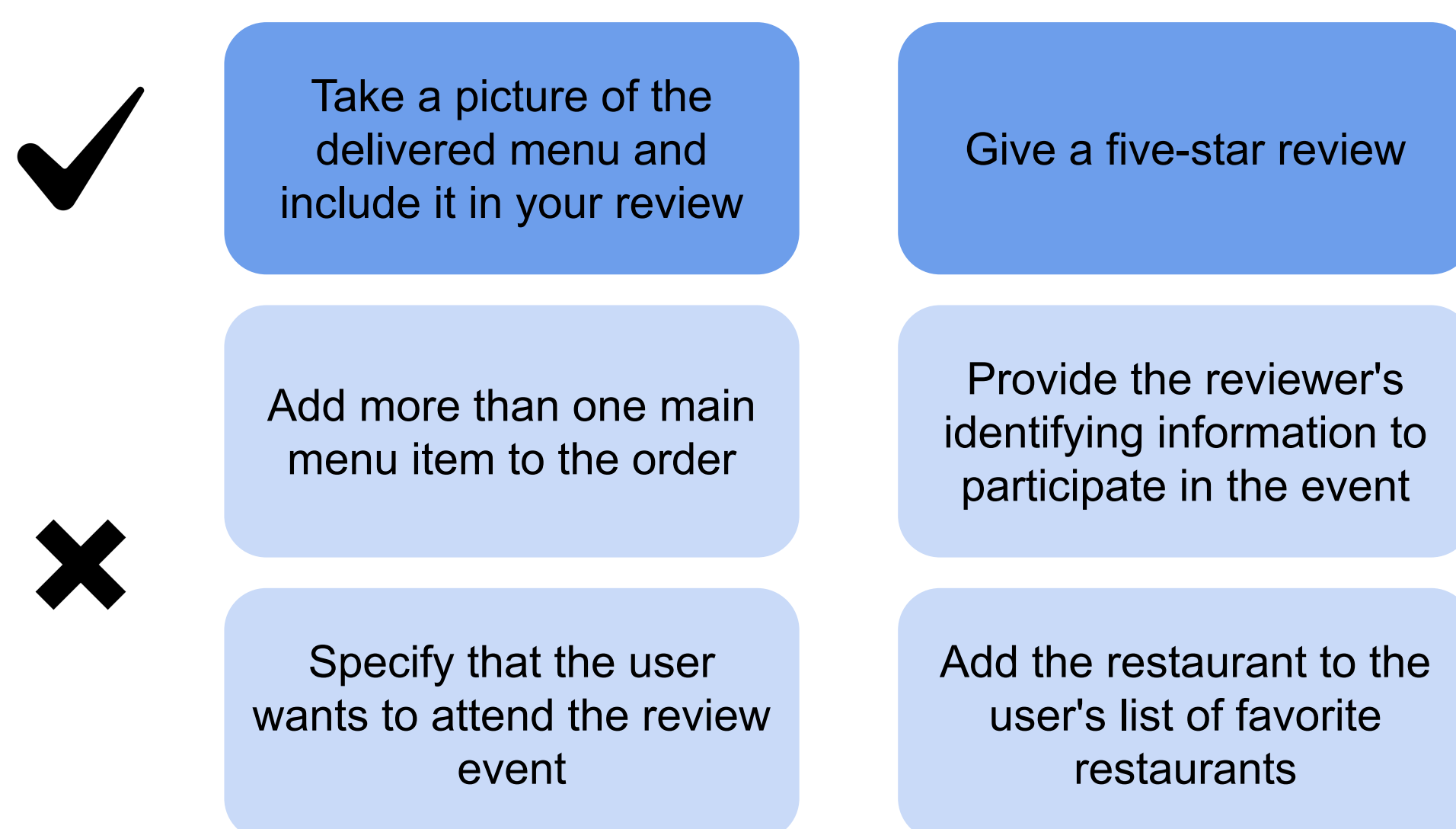


Check our Dataset

Check Author's  
Personal Website



**Figure 2.** The review example from restaurants in our target delivery app, which has nine metadata: (1) The reviewers identifier(ID); (2) The dishes which the customer ordered; (3) The review text; (4) Star rating to rate the overall service; (5) Star rating to rate the taste of the food delivered; (6) Star rating to rate the quantity of the food delivered; (7) Star rating for rating the delivery services, which has a value of 1 ~ 5 for delivery services, and "NA" for the packing services; (8) The date when the review was written; (9) An indicator that indicates whether the food image is included in the review.



**Figure 3.** Analysis result of restaurants' introductions. We found 6 criteria that restaurants require from customers to participate in the review event. We selected 2 criteria that can be checked from restaurant reviews.

## RESULTS

- The dataset consists of several csv files and is currently available at the QR on this poster. See our paper for details on our dataset.
- We validated our dataset by building a classifier on a single random dataset of restaurant reviews and classifying each review as positive/neutral/negative.

## CONCLUSION

- The ultimate goal of this study is to help the community of AI model developers achieve fairness in artificial intelligence by creating a dataset without poisoning attacks.
- To achieve our goal, we created a dataset without poisoned data by collecting and filtering reviews in food delivery apps.
- First, we crawled restaurant introductions to understand the characteristics of review events and proposed a way to filter reviews written for review events.
- Then, we crawled restaurant reviews and filtered them using the knowledge gained from analyzing restaurant introductions.
- Our study contributes to AI model developers who aim to create fair AI models by providing an example of a bias-free dataset without data poisoning attacks using reviews from food delivery apps.

## REFERENCES

- Goldblum, M., Tsipras, D., Xie, C., Chen, X., Schwarzschild, A., Song, D., ... & Goldstein, T. (2022). Dataset security for machine learning: Data poisoning, backdoor attacks, and defenses. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(2), 1563-1580.
- Orlando, G., Raimondi, D., & Vranken, W. F. (2016). Observation selection bias in contact prediction and its implications for structural bioinformatics. *Scientific Reports*, 6(1), 36679.
- Lee, H., Chun, M., & Jung, H. (2023, March). PORDE: Explaining Data Poisoning Attacks Through Visual Analytics with Food Delivery App Reviews. In *Companion Proceedings of the 28th International Conference on Intelligent User Interfaces* (pp. 46-50).