In the paper "[Saving Face: Investigating the Ethical Concerns of Facial Recognition Auditing](#)" by Raji et al., the authors present a new benchmark face dataset that provides equal representation of darker males (DM), darker females (DF), lighter males (LM), and lighter females (LF). The authors audited three facial processing technology (FPT) vendors against this dataset to assess the bias of each.

Let's imagine a use case scenario for each FPT vendor that "would result in minimal bias when in use". Of the three vendors, Microsoft performed best on the gender classification task, both in terms of overall accuracy and lighter and darker subgroups. For that reason, we could use the Microsoft gender classification service to infer the gender of customers who shop at big box stores. The Microsoft gender classifier could provide real-time gender demographic data for store owners, allowing them to understand when and, to some degree, what males and females are buying. Ultimately this information can inform store owners decision-making on product selection and presentation.

In terms of overall accuracy, the Amazon age estimator service performed better than its counterparts on the CelebSET dataset. One could argue that the difference in lighter and darker subgroups was highest for Amazon, but the overall accuracy is high enough to offset that. The Amazon accuracy between male and female subgroups was the best. We could use the age estimator service in places that sell alcohol, tobacco, and other age-restricted items. Although nothing replaces an ID, an age estimator that has high confidence that the buyer is older than some arbitrary yet safe threshold, say 40, probably doesn't have to present their ID to buy a six-pack of beer.

The Clarifai facial detection service performed better than Microsoft and Amazon in terms of overall accuracy and respectably, although not the best, in terms of subgroup accuracies. The FBI could use the Clairifai facial detection service to identify suspects in criminal cases. For such a sensitive and high-stakes use case, it would be important to have an auditing system in place to make sure that Clarifai's bias levels don't slip and its overall accuracy continues to outperform its competitors'.