Ethical Issues in the Collection of COMPAS Data

We analyzed a dataset of the results of the deployment of the COMPAS algorithm in Broward County in 2013 and 2014. One ethical issue with this data is that it is was collected when the algorithm was in use, so these results were actually displayed to judges making decisions about sentencing and bail. While we were able to learn quite a bit by analyzing the data, the deployment of the algorithm had consequences for sentencing decisions that could entirely upturn someone's life. Given some of our discoveries (which corroborate a ProPublica article analysing the same dataset cited in the introductory paper), especially the ones regarding the algorithm's reflection of the racial bias present in the criminal justice system, the cost of this data potentially included racially biased sentencing and bail decisions influenced by this algorithm. Either the algorithm hadn't been thoroughly tested, or the developers knew of the racial biases and still chose to deploy it. Both possibilities raise complex ethical questions about the fairness of the criminal justice system, and whether or not technology is doing more harm than good. The algorithm itself is just a collection of statistical techniques used to create predictions about real-world events based on past data. While our collection method simply involved downloading the dataset from the Internet, data scientists before us were making ethical decisions as they designed and then deployed COMPAS. From an educational perspective, the COMPAS algorithm furthered our own academic journey and development as statisticians. It also taught the world an incredibly important lesson about the risks of premature faith in technology. The ethically troubling question is whether or not the lives that those lessons cost were worth it, and most conventional ethical frameworks would answer no.