

# end



350,000

The growing number of developers who are part of the global ecosystem supporting Adobe Experience Cloud.



LABZ

## Fast, Private, and Fair OU Data Lab

**A**t the University of Oklahoma (OU) there are several research labs including the OU Data Lab led by Dr. Christian Grant, of which I am a member. The lab focuses on applications of fairness forensics, machine learning, and data management systems.

First up, Fairness Forensics. Dr. Kate Crawford coined the term “fairness forensics” during her keynote presentation at the 2017 Conference on Neural Information Processing Systems. At OU, researchers are working on several projects that address potential biases, such as *Detecting Simpson’s Paradox*. The phrase “Simp-

son’s paradox” is a specific type of occurrence that is a trend in a population. As an example, imagine you are a high-school student deciding which departments to apply to in a university. In your decision process, the acceptance rates at those universities may be an important factor. You start by comparing admission rates for the computer engineering and computer science departments based on last year’s data. From this comparison, you notice the computer engineering department has a better admission rate: 33 percent to 29 percent. Initially you may favor computer engineering. However, with further investigation,

you notice of those students admitted in the computer engineering department 80 percent were male and only 20 percent were female. With a deeper look into those departments’ admission rates and rejection rates, it can be noted the trends are opposite of the initial comparison. The computer science department has a higher rate of female acceptance and a lower rate of female applicants in the entire population. Thus, it can be found female admission corresponds to the subpopulation of applicants. This investigation looks at the association in the whole population that focuses on the reverses within the subpopulations defined by categorical variables. With Simpson’s Paradox, we can find surprising and interesting patterns in the data, which can be useful for companies, universities, and other entities. This is amazing work that can impact the diversity, inclusion, and programs in place at organizations.

Next up is Visual Privacy at the OU Data Lab, which explores visual privacy in the realms of IOTs, social media networks, and graph networks. The Visual Privacy project was started in 2017 and has grown into collaborations with various universities. Specifically, my research project, “VIPER,” investigates social media users posting images that contain privacy leaks in regard to themselves or someone else. These private images can include baby faces, credit cards, phone numbers, social security cards, etc. I proposed the use of supervised learning techniques train machine learning models to identify and score private images. To make this research easily integrated, we have been working to deploy a mobile application that will allow users to engage with mitigation techniques to further increase their privacy measures. This concept has been explored in IoT devices within developing smart cities. The need for visual privacy and mitigation strate-

Image courtesy of OU Data Lab



By 2022, economical shifts and the rise of emerging economies will make it difficult for any mobile provider to claim to serve a global customer base.

gies is heightened in an ever-virtual world as we continue our lives at home. Furthermore, this work can impact the exploitation of individuals and the safety protocols implemented on systems.

Finally, we take a look at Speed Labeling. To process large sets of information in exponential time, interfaces and recommendation systems are being deployed to increase the size and quality of labeled data generated per unit time. This project allows humans to enhance the performance of the labeling process by using collaborative interfaces such as touch, gesture, and brain-computer interface. The speed labeling pipeline is further enhanced by training a recommendation system, which is based on the user's activities. This addition of recommendation systems can provide data labels and best match options to help speed up labeling time. This project takes an "in-the-loop" approach by allowing humans to be actively involved in the labeling process. This project can bridge the gap between tools available for data analytics and tooling methods to avoid or investigate artifacts in data pipelines that lead to bias or unfairness.

Not only is the OU Data Lab making an impact in research, it is community focused. For the last two years, the lab has hosted Data Bite, a student-led workshop covering the basics of Python programming and machine learning. These students volunteer with local elementary schools to code robots and are involved in elementary robotic competitions. The students have also hosted tutorials, panels, and sessions at national and regional conferences. In the small, quaint state of Oklahoma, OU Data Lab is leading the way in innovation, service, and leadership.

—*Jasmine DeHart*

BACK

## The Moment In-person Instruction Ceased



We can't possibly offer our degree programs online; the quality of education would simply not be the same! That's what many institutions claimed until they were forced to move fully online in 2020. Before then, many online degree programs were associated with a stigma. Those who were unfamiliar with e-learning viewed online offerings, even from top universities, as inferior to in-person degree holders.

Notably, there was one exception: The rise of MOOCs (massive open online courses) resulted in many high-quality online courses. MOOC topics, such as machine learning, were made available without being tied to acceptance rates in the single digits. It is harder to discriminate against online education when names such as MIT and Stanford are fully online.

It should be noted many CS departments already had digital solutions in place to handle problems, such as office hours or remote login for physical lab machines, and now everyone else is catching up. There are still lingering concerns, such as outages and transitioning course material, to remain virtually engaging. But perhaps a larger unanswered problem for many platforms is student privacy. Not all platforms are transparent about student data collection, whether for that platform's use or for an instructor's use, and not all students may like features, such as emotion and facial recognition, to be utilized in the name of learning. Online university coursework isn't perfect yet, but it's getting there.

—*Daniela Zieba*