

Balancing Privacy, Quality and Unbiased Data Is Difficult, but Not Impossible

Data collection is already increasingly challenging because of emerging privacy laws. But a recent International Association of Privacy Professionals panel also noted that younger generations' privacy preferences may add even more additional hurdles.

By **Victoria Hudgins** | April 08, 2021 at 10:30 AM

The road for companies to ethically—and accurately—collect personal data will require greater transparency as more data privacy regulations and concerns arise, according to privacy experts.

On Tuesday, the International Association of Privacy Professionals (IAPP) held the virtual “The Ethics of Leveraging Data to Improve Diversity and Inclusion” panel to discuss the emerging challenges and opportunities to ethically collect and track personal data.

Macy's chief privacy officer Michael “Mac” McCullough noted that despite no federal data privacy law in the United States, current laws established well before the internet limits entities' right to control data and discriminate based on data.

“In the U.S., antidiscrimination is a qualified right vis-à-vis the government, particularly in the 14th Amendment and a patchwork of state laws,” he explained. “All of those laws have a direct disparate treatment element to them. If you blend emerging privacy laws granting rights of access, some version of consent, rights to modify and delete, we may be limiting the data underlying disparate discrimination. That's hopeful.”

Still, the enactment of the General Data Protection Regulation (GDPR), California Consumer Privacy Act (CCPA) and other privacy regulations are also blocking companies' inclusivity initiatives, he added.

McCullough noted that restrictions on the use of ethnic or other personal data will increase the use of proxy data—data that isn't explicitly linked to an individual, such as a zip code. But that could spur unintended consequences. “Proxy data by definition is not high-quality data” and can mask indirect and direct discrimination, he said.

However, another route to ethical personal data usage is to limit the collection to essential information, noted Jutta Williams, Twitter's machine learning ethics, transparency and accountability lead product manager. “My personal perspective is that fully identified information does not belong in the hands of anyone that isn't providing safety [or] health needs.”

But corporations' need for highly targeted insights requires personal data to better train artificial intelligence-based analytics models, Williams added. “We need to look at whether or not more investment into protections and more intimate data for better predictions and suggestions might be right in some industries, not all industries.”

To be sure, as companies internally grapple with ethically collecting personal data, businesses are also facing external factors pressuring their policies, McCullough noted.

“I think the day and age of employers and governments having the presumption of entitlement to data is coming to a close, especially as you look at Generation Z and Generation Alpha, both of which skew ethnic and gender categories,” he said. “Why does this matter? If ethnicity is collected for general reporting, I suspect you'll get less penetration and less high-quality data rather if you had clearly stated and measurable objectives.”