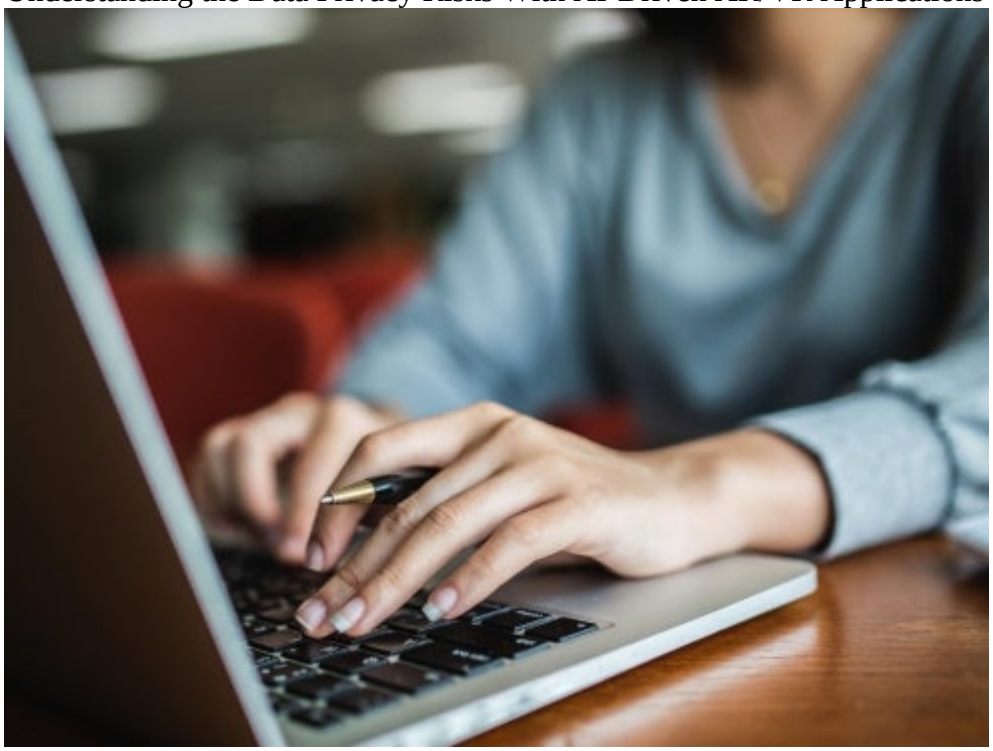
Updates
December 16, 2020
Understanding the Data Privacy Risks With AI-Driven AR/VR Applications



In the 2020 Augmented and Virtual Reality Survey conducted by Perkins Coie, Boost VC, and the XR Association, nearly three-quarters of industry leaders polled indicated that they expect immersive technologies to be mainstream within the next five years. A noteworthy number of industry leaders believe that artificial intelligence (AI) and machine learning will help drive this adoption in both consumer and business segments. The growth of immersive technologies with AI and machine learning does, however, come with risks. This article will discuss the legal, compliance, and ethical risks in the data privacy landscape when integrating machine learning functionality into immersive technology offerings.

What Are AI and Machine Learning?

Though AI and machine learning are distinct concepts, many use the terms interchangeably. AI is the broad interdisciplinary science concerning the simulation of human intelligence in machines. In practice, AI generally consists of a computer mimicking human cognitive decision processes, which may be achieved through traditional software code or via machine learning. Machine learning is a specific approach to AI focused on predictive modeling, which involves "training" algorithms on large data sets to create models that can recognize patterns and predict outcomes. These adaptive models may become more accurate over time. The survey points to AI and machine learning as separate factors accelerating adoption of immersive technology, but because many of the current implementations of AI technology are based on machine learning, this article focuses primarily on the ways machine learning—and not the broader concept of AI—is helping to overcome what survey respondents identified as the key barriers to adoption.

What Data Privacy Concerns Are Raised by the Use of Machine Learning Models in Immersive Technology Applications?

Machine learning is accelerating adoption of immersive technologies in several important ways, including through advancements in user experience, enhancements to content quality, and greater adaptability to user input in a variety of situations. Generating an advanced and useful machine learning model is only made possible by the use of significant amounts of data, often consisting of user data (including personally identifiable information (PII)), system data, and analytics data. Both the collection and the use of this data may raise potential legal, compliance, and ethical risks that companies may need to address in order to avoid liability or fines. This is consistent with survey responses identifying data privacy and security as one of the top legal concerns. Increasingly, data collection and processing is being regulated by privacy laws, including the General Data Protection Regulation (GDPR) in the European Union (EU) and, in the United States, recently enacted state laws such as the California Consumer Privacy Act (CCPA), the California Privacy Rights Act (CPRA), and the Illinois Biometric Information Privacy Act (BIPA). These data privacy regimes raise concerns about the development and deployment of immersive technologies that utilize machine learning. Further, algorithmic or data set bias presents ethical risks that companies should also consider.

Personal Information

Certain privacy regulations of foreign countries and several U.S. states may apply when immersive technology applications gather or leverage PII, such as personal identifiers, biometric data, search and purchase histories, geolocation data, and other inferences. Augmented reality (AR) applications, in particular, may collect significant amounts of PII as they register and map the environment. This information may then be processed by a machine learning algorithm in order to refine the model powering an application's functionality. In other instances, companies may purchase or leverage third-party data sets containing PII to train their algorithms. In either case, use of this PII has key implications under the GDPR and U.S. state privacy regulations.

GDPR

The GDPR regulates the collection, use, retention, and disclosure of PII of EU residents. Penalties for noncompliance with GDPR are significant. Fines can be as high as the greater of €20 million or up to 4% of the annual worldwide turnover in the preceding financial year of the offending company. These penalties emphasize the importance of complying with GDPR obligations when integrating machine learning into immersive technologies. The GDPR affects the use of machine learning in at least three key ways: (1) by limiting the collection and use of data; (2) by restricting automated decision-making in certain scenarios; and (3) by imposing obligations on companies to disclose their PII collection and use practices to users of an application and to insert certain required contractual protections into commercial agreements concerning the processing of EU residents' PII.

First, Article 6 of the GDPR restricts companies' ability to collect and process PII from EU residents unless the company obtains freely given, specific, informed, and unambiguous consent from the data subject or otherwise has a legitimate business purpose for collecting that PII and processes the PII only for that limited purpose. To remain compliant, companies must document and demonstrate that both they and their licensors have either complied with the GDPR's requirements for the processing of PII, including obtaining any required consent, or properly aggregated and anonymized PII in their data sets. If a company is collecting and using PII for the

purpose of training machine learning algorithms and relying on a consent mechanism, that specific purpose must be made clear to the data subject to help ensure valid consent.

Second, Article 22 of the GDPR provides EU residents the right to object to automated decision-making (e.g., through AI or machine learning), such as profiling, and requires a company to inform EU residents how they can object to or request human review of the automated decision-making. For any decision that is based solely on automated processing where the decision is either not authorized by law or not necessary for the performance of a contract between the individual and the data controller (i.e., the company), the individual's explicit consent is required. When an immersive technology application uses machine learning to make an automated decision that "significantly affects" an individual's rights, for example, an AR training application that takes data from an employee's training evaluation and runs it through a machine learning model to sort the employee into a job function within the company or AR headsets with integrated facial recognition software used by law enforcement to make decisions in the field, companies should ensure the application is designed and developed with sufficient levels of human review. This human review will help ensure that the algorithms and models are operating as intended, ensure the accuracy of results, and prevent automation bias.

Third, the GDPR requires that companies make certain disclosures to users about their data practices and impose certain restrictions on their subcontractors who process PII on the company's behalf (or on whose behalf the company processes PII). Companies must create publicly available privacy notices with adequate and accurate disclosures regarding their practices when handling and storing PII and ensure that any agreements with contractors involving the processing of EU residents' PII include the clauses required by the GDPR.

CCPA and CPRA

Multiple U.S. states have passed consumer protection laws that offer some level of protection for a company's use of PII, including laws covering data breach notification and those specific to health information. By contrast, California's CCPA is the most comprehensive state legislation in the United States to date. The CCPA broadly defines the meaning of PII and imposes various notice, verification, and opt-out requirements on companies that collect the PII of California residents. The CPRA, which heightens the CCPA's privacy requirements, passed the California ballot on November 3, 2020, and will become effective January 1, 2023.

Noncompliance with the CCPA carries stiff penalties. Violators of the CCPA may be liable for civil penalties up to $7,500 per intentional violation. If an immersive technology company fails to implement and maintain "reasonable security" to protect PII, the company may also be subject to damages up to the greater of $750 per consumer per incident or actual damages. Immersive technologies may collect several types of PII subject to the CCPA, such as geolocation data to inform a user's visual experience, facial or voice data for user identification capabilities, or an inference made by a machine learning model that could be reasonably associated with a particular California consumer. The CPRA further expands the definition of "sensitive" PII to include a consumer's race or ethnicity, nonpublic communications, and biometric data that uniquely identifies a California consumer.

While imposing disclosure obligations similar to the GDPR, the CCPA also requires that companies give users the right to opt out of the sale of their PII. In addition to the generally understood definition of sale, combining PII with third-party data to augment certain types of data sets may also be considered a "sale" of PII under the CCPA, which would require notifying users and giving them the opportunity to decline that use of their PII. Data set augmentation is a likely use case since significant value can be gained by improving the data sets used to train the machine learning algorithms. The CCPA also requires that companies provide a "just-in-time" notice before collecting PII if the collection is for a purpose that the customer may not reasonably expect (such as geolocation information to feed into a machine learning model for optimizing AR mapping). Companies will need to conduct an analysis to determine whether their collection or use of PII constitutes a "sale" governed by

the CCPA and whether such collection or use would be "reasonably expected" by the user.

In addition to the obligations under the CCPA, the CPRA will require companies to give their users the right to limit the use of their sensitive PII, including the ability to opt out of, or obtain information about, automated decision-making technology and to correct inaccurate PII. While preparing for CPRA compliance, a company should determine what, if any, sensitive PII its immersive technology application collects, uses, and shares, and ensure that facilitating a user's right to limit the company's use of that user's sensitive PII will not impair the functionality of the company's technologies.

BIPA

Other state privacy laws, including those of Illinois, Washington, Texas, New York, and Arkansas, regulate biometric information, which includes fingerprints and face, hand, retina, or ear features. These privacy laws require the individual's informed consent before a company can obtain or disclose that individual's biometric information. Currently, Illinois's BIPA is the only state biometric law that expressly offers a private right of action. BIPA violations carry potentially significant liability from class action lawsuits—ranging from $1,000 to $5,000 per violation. This implicates AR applications in particular because they have a high likelihood of collecting biometric data through their use.

To navigate the risks presented by regulations such as the GDPR, CCPA, and BIPA, companies should familiarize themselves with the types of PII their immersive technology applications collect and how such PII is used (including for the purposes of training machine learning algorithms), how the machine learning models in their applications work, and whether any inferences that are generated from those models rise to the level of PII that is subject to data privacy law.

Bias in Machine Learning Algorithms or Data Sets

Both data sets and the machine learning algorithms used to generate machine learning models can contain bias, whether intentional or unintentional, that can adversely impact the effectiveness of the application in which the models are embedded. This creates ethical concerns around the application's use and can lead to potential discrimination claims. The demographics of users, often arising from institutional barriers to access, can skew the data in a machine learning training data set. An unrepresentative data set can cause a machine learning model trained on that data to form inaccurate conclusions or display prejudice in its pattern identification, with this bias most often harming marginalized groups. In addition to the negative ethical implications, this can expose companies to liability for discrimination, particularly if the immersive technology being powered by machine learning is used to make important decisions regarding an individual's rights. For example, respondents to the survey identified workflow management as an immersive technology application that companies are likely to focus on over the next 12 months. If bias in the machine learning model powering that application systematically denies assignments and opportunities to a certain group of employees due to that bias, this could give rise to a claim of workplace discrimination.

Machines and algorithms can produce results that exhibit racist bias or have another discriminatory impact. As Joy Buolamwini, a computer scientist and researcher at Massachusetts Institute of Technology, explained, "If computers are trained on data sets of photos in which people of color were absent or underrepresented, the system may fail to recognize people of color as people." Buolamwini's research revealed that some facial recognition software from tech giants could not identify women of color. To prevent inadvertent bias, both for the sake of the user experience and to avoid potential liability, machine learning algorithms should be trained with a diverse data set and vetted for biased decision-making.

Authors

Explore more in

Emerging Companies & Venture Capital Law | Technology Transactions & Privacy Law | Corporate Law | Privacy & Security | Artificial Intelligence & Machine Learning | Immersive Technology

Related insights

Update

California Court of Appeal Casts Doubt on Legality of Municipality's Voter ID Law

Update

February Tip of the Month: Federal Court Issues Nationwide Injunction Against Trump Executive Orders on DEI Initiatives