Published on *The National Law Review* https://natlawreview.com

Five Key Considerations to Developing Defensible AI

Article By:

Kevin Heaphy

The possibilities for artificial intelligence applications in business seem unlimited. Algorithms are being developed, and in some cases are already implemented, to handle customer inquiries, process loan applications, train employees, analyze market trends, evaluate job applications, hyper-personalize customer experiences, and undertake many other tasks. Companies in numerous industries are harnessing the value of data to make business processes more efficient and improve their products and services. In a recent survey published by Forbes, 91% of enterprises said they expect AI to deliver new business growth in the next five years.[1]

The rapidly expanding implementation of artificial intelligence in society means that, inevitably, lawyers will be faced with putting artificial intelligence on the witness stand. Whether it is to explain an autonomous vehicle’s lane change, the intent of an autonomous smart contract, or the basis of an autonomous hiring decision, lawyers will soon be faced with defending the actions of faceless algorithms. What should companies implementing AI solutions do now to ensure that they can effectively defend their use of AI when the machine becomes the witness? This article offers some thoughts on developing AI in a defensible way in order to be prepared should the AI system itself become the focal point of a lawsuit.

Pay Attention to the Data Used to Train Your AI

AI training requires a large volume of data to learn the operations that can lead to improved efficiency and generate revenue. Companies looking to implement AI programs must first secure the rights and access to the data. Further, in order to train AI, the data must be selected, cleaned, and prepared to ensure the best results. In fact, many companies are sprouting up that specialize in simply cleaning data to support the development of AI. In the early stages of training AI, decisions regarding what data should be included or excluded, and how the data should be labeled, are being made by programmers. This poses a problem for enterprises that may end up defending decisions made by algorithms without a clear understanding of how they were trained.

A key step in planning for defensible AI is being mindful of the people who will be responsible for making decisions regarding the data used to train machines. In particular, these individuals must understand the potential for wrongfully influencing data sets with seemingly benign decisions – for example, a geographic limitation that results in racially biased decision making. Further, the person or team assuming this responsibility should document the data used to train the machine, including which specific categories were used and how the data was prepared. Documentation should also memorialize the decisions they made during training, their reasoning for them, and any observations made after training is complete. Having this type of information, and the individuals who understand it, available as a source of evidence is a significant component of a defensible AI program.

Understand the Black Box

AI systems often take the form of a black box, with an input, an output, and little understanding of what happens in between. But an organization must be able to explain *why* a decision was made by its AI system. Imagine an AI system that consistently denies loans to applicants of a certain race, or a job application algorithm that predominantly selects white males for higher-paying positions. Companies employing AI will need to be able to explain the decision-making process in order to defend their systems. Best practices include documenting the machine language, or the decision-making steps, in an understandable format in the event you need to justify it later. Ideally, your programmers should be capable of both understanding your AI systems and explaining them easily. Where possible, AI systems should also be programmed to produce an explanation of their reasoning along with their results and quality control policies should incorporate a review of explanations for potential issues. Taking these steps to develop explainable AI will avoid situations where a company is unable to describe how its algorithms operate when faced with an allegation in a lawsuit.

Develop AI in an Ethically Responsible Way

Many organizations recognize the ethical development of AI as a priority for their business development. In fact, various organizations, such as the International Technology Law Association (ITechLaw), are increasingly publishing guides and frameworks for the responsible development of AI with a focus on ethics.[2] Companies should consider framing the development of AI in ethical and responsible procedures from the start. The main focus on developing AI in an ethically responsible manner is to focus on things like accountability, transparency, and reliability. The advantage in doing so is twofold: the company is less likely to develop an AI system that triggers a litigation risk and, in the event it does, the company will be able to effectively demonstrate the ethical and responsible approach of its development process.

Be Mindful of Privacy Regulations

With the GDPR and the CCPA going into effect, and similar acts being considered or approved in New York, Nevada, and other states, most companies recognize the importance of taking measures to protect the privacy of data they collect. As AI systems must typically digest a large volume of data to become reliable, companies must pay special attention to the type of data used and how it may violate the web of privacy regulations taking effect. For example, if the data includes personally identifying information (such as names, dates of birth, addresses, etc.) the company may have a duty to disclose how that data is being used to the consumer. Likewise, companies using this data will have to be prepared to remove the data from the database, or anonymize it, if consumers exercise their rights to have their data deleted. Companies will also need to ensure that the process of pseudonymization and anonymization—if utilized—is sufficient to render the data personally unidentifiable. These considerations, and many more, should be addressed early on in the development of AI programs to ensure that the use of AI does not trigger a significant liability.

Always Look to Improve Results

The reality is that customer complaints are not likely to disappear as long as there are customers. As AI systems increasingly interface with customers, complaints will be the early red-flags to companies that a risk of potential litigation is looming. For example, a customer complaining that a real estate algorithm is not showing all available properties may have identified a potential bias in the AI. Taking complaints seriously, and treating them as an opportunity to improve your results, is critical. Companies should have quality control and troubleshooting policies in place to address such concerns and follow those policies to improve AI systems and reduce their overall risk.

Conclusion

As AI systems become more prominent in delivering products and services, companies will not be able to avoid liability by being ignorant of their own AI. Companies must take steps to understand their own AI, and the data it uses, to maximize its defensibility.

References:

[1] https://www.forbes.com/sites/louiscolumbus/2018/09/29/91-of-enterprises-expect-ai-to-deliver-new-business-growth-by-2023/#5261e92e6959.
[2] Available online at https://www.itechlaw.org/ResponsibleAI.

National Law Review, Volume IX, Number 270

Source URL:https://natlawreview.com/article/five-key-considerations-to-developing-defensible-ai