AN A.S. PRATT PUBLICATION
OCTOBER 2020
VOL. 6 • NO. 8

PRATT'S PRIVACY & CYBERSECURITY LAW REPORT

LexisNexis

Pratt's Privacy & Cybersecurity Law Report

VOLUME 6 **NUMBER 8** **OCTOBER 2020**

LexisNexis®

QUESTIONS ABOUT THIS PUBLICATION?

For questions about the **Editorial Content** appearing in these volumes or reprint permission, please contact:
Deneil C. Targowski at 908-673-3380
Email: ..... Deneil.C.Targowski@lexisnexis.com
For assistance with replacement pages, shipments, billing or other customer service matters, please call:

Customer Services Department at ..... (800) 833-9844
Outside the United States and Canada, please call ..... (518) 487-3385
Fax Number ..... (800) 828-8341
Customer Service Web site ..... http://www.lexisnexis.com/custserv/
For information on other Matthew Bender publications, please call

Your account manager or ..... (800) 223-1940
Outside the United States and Canada, please call ..... (937) 247-0293

ISBN: 978-1-6328-3362-4 (print)
ISBN: 978-1-6328-3363-1 (eBook)

ISSN: 2380-4785 (Print)
ISSN: 2380-4823 (Online)

Cite this publication as:
[author name], [*article title*], [vol. no.] PRATT'S PRIVACY &CYBERSECURITY LAW REPORT [page number] (LexisNexis A.S. Pratt);
Laura Clark Fey and Jeff Johnson, *Shielding Personal Information in eDiscovery*, [5] PRATT'S PRIVACY & CYBERSECURITY LAW REPORT [245] (LexisNexis A.S. Pratt)

An A.S. Pratt Publication
Editorial

Editorial Offices
630 Central Ave., New Providence, NJ 07974 (908) 464-6800
201 Mission St., San Francisco, CA 94105-1831 (415) 908-3200
www.lexisnexis.com

MATTHEW BENDER

(2020–Pub. 4939)

Editor-in-Chief, Editor & Board of Editors

Pratt's Privacy & Cybersecurity Law Report is published nine times a year by Matthew Bender & Company, Inc. Periodicals Postage Paid at Washington, D.C., and at additional mailing offices. For customer support, please contact LexisNexis Matthew Bender, 1275 Broadway, Albany, NY 12204 or e-mail Customer.Support@lexisnexis.com. Direct any editorial inquires and send any material for publication to Steven A. Meyerowitz, Editor-in-Chief, Meyerowitz Communications Inc., 26910 Grand Central Parkway Suite 18R, Floral Park, New York 11005, smeyerowitz@meyerowitzcommunications.com, 646.539.8300. Material for publication is welcomed—articles, decisions, or other items of interest to lawyers and law firms, in-house counsel, government lawyers, senior business executives, and anyone interested in privacy and cybersecurity related issues and legal developments. This publication is designed to be accurate and authoritative, but neither the publisher nor the authors are rendering legal, accounting, or other professional services in this publication. If legal or other expert advice is desired, retain the services of an appropriate professional. The articles and columns reflect only the present considerations and views of the authors and do not necessarily reflect those of the firms or organizations with which they are affiliated, any of the former or present clients of the authors or their firms or organizations, or the editors or publisher.

POSTMASTER: Send address changes to *Pratt's Privacy & Cybersecurity Law Report*, LexisNexis Matthew Bender, 630 Central Ave., New Providence, NJ 07974.

Training a Machine Learning Model Using Customer Proprietary Data: Navigating Key IP and Data Protection Considerations

By Brittany Bacon, Tyler Maddry, and Anna Pateraki *

In this article, the authors examine intellectual property ownership and licenses related to the machine learning process, and how companies can ensure that data protection and security-related obligations are addressed and potential risks are mitigated.

The use of machine learning ("ML") models to process proprietary data is becoming increasingly common as companies recognize the potential benefits that ML can provide. Many IT vendors offer ML services that can generate valuable insights derived from their customer's proprietary data and know-how. For companies that have not yet established their own ML expertise in-house, these services can offer significant business advantages.

However, there may be cases where one party owns the ML model, another party has the business expertise, and a third party owns the data. In such cases, significant intellectual property ("IP") and data protection and security risks may arise.

Data protection regulations in the United States, Europe, and beyond dictate which types of data processing are permitted and under what conditions, as well as what agreements should be executed between the parties to delegate their respective data-related roles and responsibilities.

Additionally, to protect a business's IP, it is necessary to understand the different elements of the ML process, such as the training data, methods of training, types of ML models, results of the ML, and the different types of IP that protect these elements. All of these considerations go into implementing processes and contractual terms to specify IP ownership and licenses related to the ML process, and to ensure that data protection and security-related obligations are addressed and potential risks are mitigated.

---

* Brittany Bacon is a partner in Hunton Andrews Kurth's global privacy and cybersecurity practice, resident in the firm's New York office. Tyler Maddry is a partner in the firm's intellectual property practice, resident in the firm's Washington, D.C., office. Anna Pateraki is a senior associate in the firm's global privacy and cybersecurity practice, resident in the firm's Brussels office. The authors may be contacted at bbacon@huntonak.com, tmaddry@huntonak.com, and apateraki@huntona.com, respectively.

BACKGROUND

Artificial Intelligence ("AI") is an umbrella term for a range of technologies that attempt to perform human-like cognitive processes. Things that humans have traditionally done by thinking and reasoning are increasingly being done by, or with the help of, machines exhibiting "intelligent" behaviors, such as speech and facial recognition, interpretation of medical images, providing personal recommendations for products, verifying identity, and detecting fraudulent behavior.

Machine learning ("ML") is a subset of AI that involves methods of modifying or improving the execution of an algorithm coded in a computer program (the "ML model") based on the use of training data. ML enables a computer system to learn and adapt without being explicitly programmed with predetermined rules.

In practice, there often are multiple parties involved in successfully applying an ML model to a particular problem or opportunity.

For example, it is common to have one party who has expertise in a particular industry (e.g., a medical imaging device manufacturer), a second party who has expertise in implementing ML models (e.g., IT vendor), and a group of individuals from whom valuable data is collected to train the ML model (e.g., patients). There also may be experts who are needed to label the training data (e.g., a pathologist to label a biopsy image as carcinogenic or benign).

The rules for who can do what with the data and IP are determined to a large extent by applicable data protection laws and IP laws, as well as agreements between the various parties.

Many data protection laws, such as the Health Insurance Portability and Accountability Act ("HIPAA"), the Gramm-Leach-Bliley Act ("GLB") and the California Consumer Privacy Act of 2018 ("CCPA") in the U.S., as well as the General Data Protection Regulation ("GDPR")in the European Union ("EU"), define the rights of the individual data subjects and the responsibilities of the entities that access or otherwise process the individuals' personal information. IP laws define the rights each party has in certain assets such as software, proprietary information, know-how, proprietary processes, inventions, derived data, and other results. Contracts can modify the IP rights between parties, including granting permissions to use proprietary data under specified conditions.

Naturally, most companies that invest in building an ML model are looking for a return on their investment. From a financial perspective, such companies focus on using the IP laws and related IP contract terms, such as IP assignments and license grants, to maximize their control over the ML model and associated input and results. Data protection laws can run counter to these objectives by imposing an array of requirements and restrictions on the processing of various types of data, particularly to the extent they

include personal information. The interplay between these competing considerations can lead to interesting results, especially when a number of different parties have a stake in the outcome.

OVERVIEW OF MACHINE LEARNING

The defining characteristic of a machine learning model that distinguishes it from a typical, fixed algorithm running in a computer program is that the ML model uses data to improve its performance. That data is generally referred to as "training data" because it is used to train and improve the ML model. It does this by adjusting certain parameters used within the ML model. For example, the ML model may be designed to analyze a medical image, such as a microscope slide image of a tissue sample. The ML model is trained to determine whether the slide image shows the presence of a particular disease.

The method of training may involve labeling the image, such as a pathologist viewing the image, determining whether it shows the disease, and adding labeling information to the image (e.g., indicating that the disease exists in certain regions of the image). Labeled images (training data) are then used to iteratively update the parameters of the ML model.

For example, a slide image is used as an input to the ML model, and the ML model outputs its results, such as whether the slide image shows the disease and a factor indicating how confident the ML model is in its conclusion. The results output by the ML model are compared to labeling information provided by the pathologist, and the parameters that are used in the ML model are adjusted to improve the fit between the results calculated by the ML model and the labeling information.

As more and more training data is run through the ML model, the parameters used in the model continue to be refined, and the accuracy of the ML model improves (i.e., it "learns") over time. Eventually, the training data tunes the parameters of the ML model so well that it can be used to characterize unlabeled data (a new, unlabeled slide image). ML models can provide significant advantages over traditional analytical methods that rely only on algorithms, rather than data, especially when it is difficult or impossible to create those algorithms.

Anyone who has dipped their toe into the world of ML knows there are many types of ML models and training methods, but the fundamental concept of using training data to adjust and improve the execution of the ML model is generally applicable across the different variations.

IP OWNERSHIP AND USE RIGHTS

From an IP perspective, there are different types of IP rights, such as copyrights, patents, trade secrets and database rights that may be used to protect the various components of the ML model and associated data.

The ML model itself comprises software and certain parameters that control and adjust its execution. The source code of the software may embody proprietary algorithms that constitute trade secrets of the software owner. Indeed, for developers of proprietary software, the source code is often a very closely held trade secret. Software may also be protected under copyright law as a literary work.

In addition, certain components of the software may qualify for patent protection if they embody an inventive process that is not obvious in view of what is in the public domain (prior art). The parameters used to control the operation of the software, and the algorithms used in the ML model, may also embody valuable information that the software developer considers a trade secret.

The training data, particularly if it is labeled, may constitute a trade secret if it has been maintained as confidential using appropriate safeguards. It also may qualify for protection under database rights laws in Europe and potentially the United States.

Methods of training the ML model may qualify as trade secrets. For example, the ML model owner typically has developed valuable know-how in deciding what training methods work best for certain ML model types. Moreover, a business owner with expertise in a particular industry may use its business trade secrets in training the ML model, such as what data elements are most significant for their predictive value, what data elements are irrelevant, and how to label the training data. In fact, incorporating these types of business trade secrets into an IT vendor's ML model can be one of the more sensitive issues in a negotiation.

Of course, the reason for building and training the ML model is so that it can output valuable results and insights. To the end, the U.S. Patent and Trademark Office ("USPTO") recently entertained the question as to whether a computer program could be an inventor.[1] Although the answer was no,[2] the output from an ML model can qualify as a valuable trade secret or proprietary information.

Given the various rights available in the ML context, parties are well advised to analyze and address the question of which party owns the IP rights in the different elements of the ML model at the outset. An IT vendor typically will own the ML model software and associated trade secrets, copyrights and patents. It also will own any trade secrets in its training methods. A business owner will own its business trade secrets in methods of training the ML model and also may have certain rights in the training data. Although the default rules on IP ownership provide the starting point, the more

[1] *See* USPTO Request for Comments on Patenting Artificial Intelligence Inventions, Fed. Reg. Vol. 84, No. 166 p. 44889 (Aug. 27, 2019).

[2] *See* In re Application No. 16/524,350, FlashPoint IP Ltd., Decision on Petition (2020), *available at* https://www.uspto.gov/initiatives/artificial-intelligence.

important considerations relate to the contract terms that assign ownership of IP rights and/or grant licenses or permissions to use the components of the ML model, as will be discussed below.

DATA PROTECTION LAW CONSIDERATIONS

As will often be the case, to the extent an ML model involves the processing of personal information (which typically refers to information that relates to an identified or identifiable individual), the relevant parties will need to consider the legal and regulatory requirements and restrictions that apply to the collection, use, and disclosure of such information.

Which requirements and restrictions come into play depend on which data protection law(s) apply from a jurisdictional perspective, an issue that typically takes into consideration where the respective parties are established or do business, and/or the residency of individuals whose personal information is being used.

In the U.S., data protection regulation historically has comprised a patchwork quilt of laws that primarily embrace notice and choice as guiding principles. With the recent enactment of the California Consumer Privacy Act of 2018 ("CCPA"), the U.S. legal regime is shifting to one that offers individuals certain rights with respect to their data (i.e., access, deletion, and opt out of sale), moving away from the notion that businesses that collect the data are "owners" of such information with the autonomy to use the data indefinitely and without question as long as appropriate notice and choice were offered at the outset.

In the EU, where privacy is a fundamental human right, the GDPR represents an omnibus data protection law that establishes more rigid requirements and restrictions and grants an array of data protection rights to data subjects without relying on a distinct concept of data ownership. Many countries outside of the EU have adopted a GDPR-style of regulation that is further informed by their historical cultural experiences.

Beyond the data protection rights afforded to individuals and various interpretations of data ownership under applicable law, data protection regulators (such as the U.S. Federal Trade Commission) have urged companies to carefully consider and combat the unintended risks that may arise with ML, such as discrimination, bias, and automated decision-making that may impact the rights and freedoms of individuals.

Additionally, many data protection laws and regulations impose specific obligations on the parties involved in the processing of personal information, and require that the contractual terms reflect the respective roles and responsibilities.

While relevant data protection laws prescribe minimum contractual terms to govern the roles and responsibilities of the parties in the processing of personal information, additional contractual terms may be required or appropriate in a specific case to manage

data protection risks in the ML context. Before drafting such additional contractual terms for ML purposes, it is important to understand the relevant requirements, restrictions and data protection risks at all stages of the ML lifecycle, including in the training and testing phase whereby a model is created, and in the deployment phase where the model is applied to a specific case. We provide below an overview of the key data protection issues in the ML context that can be addressed through additional contractual terms, as appropriate.

In the training and testing phase, it is important to take steps to ensure the fairness and accuracy of the ML model, and prevent unwanted bias. In addition, the ML model should apply appropriate privacy-enhancing technologies, including technical and organization measures, to mitigate privacy and security risks to personal information. Such measures may include pseudonymization and encryption where possible, as well as privacy, security and ethics-by-design measures that shape the lifecycle of the data processing.

To that end, when engaging ML models developed by an IT vendor, as part of the vendor due diligence process, a business owner should inquire about the compliance posture of the vendor's ML processes and assess the vendor's ability to implement these risk-mitigating measures.

In the deployment phase, as a data controller (or equivalent under applicable law), the business owner is responsible for the laws applicable to the processing of personal information using ML technology (including the sharing of customer data with an IT vendor that offers ML services), and the continued training and improving of the ML model for the business owner's purposes. Subject to applicable law, the business owner should include disclosures relevant to the ML purposes in its privacy notices, ensure that it relies on an appropriate legal basis for the processing of personal information, and have processes in place to comply with rights requests from data subjects (e.g., right of access to or deletion of personal information). Where applicable law restricts automated decision-making that could result in significant consequences for data subjects, business owners or relevant IT vendors should incorporate human review into such decisions and recommendations. Some IT vendors, for example, have begun to offer ML services that enable human intervention in response to a demand for such tools to facilitate compliance with applicable privacy laws.

In practice, there is a tension between complying with applicable data protection requirements and principles and maximizing the effectiveness of ML tools. For example, the general principle of data minimization under EU and other data protection laws (i.e., use the minimum amount of data required to achieve a certain purpose) may well conflict with the principle of accuracy in the ML context.

On the one hand, the more data used for ML training, the more accurate the ML process is likely to be.

On the other hand, the more data that is processed to pursue accuracy, the greater the potential risk to privacy as larger volumes of personal information are being processed. This tension can be addressed by conducting a relevant privacy impact or risk assessment that seeks to balance the benefits of ML technology against the trade-offs and potential negative consequences for data subjects, helping to identify risk mitigation measures at an early stage of using ML technology.

From a governance perspective, both business owners and IT vendors should also consider leveraging existing structures of their data protection compliance program to address ML issues in a thoughtful manner and to implement accountability measures for their use of ML. These measures may include:

- Ensuring appropriate leadership, governance, and oversight of ML issues (e.g., collaboration between senior management, privacy team/Data Protection Officer ("DPO"), other data protection experts vested in AI issues, data scientists or privacy engineers), and
- Taking steps to ensure accountability and documentation of the relevant ML processes through records of data processing activities, training materials, and internal policies and procedures, as applicable.

DRAFTING CONTRACT TERMS

IP Terms

It is common to see various parties in an IP contract negotiation all insist on exclusive IP ownership which, of course, is not possible. One way to move beyond the impasse is to focus less on abstract IP rights and more on what each party actually needs to do in its business. If there are questions about which party should own IP that is generated or used in a collaboration, the factors to consider usually include:

- Which party or parties created the IP (e.g., software, know-how, proprietary processes, output);
- Whether the developed IP is in one party's core technology area; and
- Which party will need to enforce the IP rights against third parties.

Answers to these questions usually provide a good indication as to which party should own the developed IP rights.

It is also essential to identify, define and understand all of the components of the ML model, including the origin and elements of the input training data, the proprietary business information that will be used to train the ML model, the IT vendor's methods of training the model, the ML model itself, improvements to the ML model, and the output from the ML model. An investment up front in learning the technical facts

makes the contract drafting and negotiation much more efficient than negotiating IP rights in the abstract. It is also worthwhile to understand each party's business model and how it expects to make a financial return on the collaboration, as this will illuminate the motivations of the other party to gain access to various components of the ML model and related data.

Although it is important to assign IP ownership in a way that gives your company the most control, the more useful way of reaching agreement with the other party is through appropriate licensing terms. Licensing of IP rights allows the parties to define a very detailed set of rights and restrictions for different components of the IP that can be tailored to the business.

In drafting these IP rights and restrictions, it is necessary to first define each component of IP and technology that will be licensed. Separate license grants can be drafted as necessary to account for different permitted uses of the various IP and technology components. Each license grant should be defined with a number of characteristics that typically include:

- The proper licensee or licensees;
- Exclusive versus nonexclusive rights;
- Permitted activities;
- Field of use restrictions;
- Term/duration;
- Revocability;
- Transferability;
- Territorial restrictions; and
- The right to sublicense or engage third parties.

In addition to the license grant, it will usually be necessary to draft other restrictions such as confidentiality obligations and explicit use restrictions. In drafting these terms, it is important to keep in mind that a confidentiality obligation does not by itself preclude use of information (i.e., a receiving party can use confidential information for many purposes while keeping it confidential). Defining certain data sets as proprietary or confidential with associated obligations also enhances protection of the data, particularly if the data does not qualify as a trade secret.

From the perspective of a business owner using an IT vendor to provide ML services, it may be especially important to consider what the IT vendor is permitted to do with the input and output.

In particular, the process of choosing and configuring the training data and the methods of configuring and training the ML model may involve collaboration in which the business owner discloses valuable trade secrets of its business to the IT vendor (i.e., to the IT vendor's employees not just to the ML model). The IT vendor will want the freedom to engage similar clients in the same industry, but the business owner cannot let the IT vendor share its trade secrets with a competitor via the IT vendor's employees or trained ML model.

Hence, the business owner needs to carefully define restrictions on use of its know-how and other IP, while the IT vendor will want to preserve as much freedom as possible to engage clients in the same or a similar industry.

Data Protection Terms

Business owners with global operations must comply with various applicable legal regimes when engaging an IT vendor to train an ML model using the business owner's customer data. Data protection laws in the U.S. and the EU, for example, impose contractual requirements relating to data sharing, data sourcing and processing that the IT vendor will perform on behalf of the business owner.

In the United States, various state and sector-specific laws require businesses to enter into written agreements with service providers (similar to data processors in the EU) that limit the service provider's ability to process the data for any purpose other than to perform the services and to employ reasonable safeguards to protect the data. A key consideration when entering into a contract with an IT vendor is to ensure that the vendor's access to and use of such data does not run afoul of representations the business owner has made to data subjects whose personal information is being processed in connection with the ML model.

Apart from this purpose and use limitation, the data protection-related contractual provisions tend to focus on allocating responsibilities and potential liabilities, and it is important for business owners and IT vendors to consider their business objectives and risk tolerance in negotiating such terms.

In the EU, under the GDPR, the responsibilities of the parties will depend on their data protection role as a data controller or processor of personal information.

A data controller is an entity that determines both the "purposes" and "means" of the data processing, meaning both the "why" and "how" the personal information is processed to achieve a certain business purpose. This includes making autonomous decisions about core elements of the data processing, such as which data elements will be collected or processed, whether the processing should be changed, where and how the data will be stored, if a vendor will be engaged to assist with the data processing, and for how long the data will be retained to achieve a certain purpose.

A data processor is an entity that processes personal information on behalf and under the instructions of the data controller under the relevant contract.

With respect to responsibilities, a data controller is directly subject to the full spectrum of the GDPR's requirements, in particular the requirements relating to data subjects (e.g., to provide notice, obtain consent where required, respond to data subject rights requests, report data breaches to regulators and affected individuals, etc.). A data processor, on the other hand, processes the data in the back-end and is subject to more limited legal obligations under the GDPR, such as data security requirements and accountability obligations (e.g., to appoint a DPO, if applicable, and maintain records of data processing activities), along with the contractual commitments that restrict the processor's ability to process the personal information it receives from, or collects on behalf of, the data controller.

A data processor must not process the personal information for the processor's own purposes, or it becomes a data controller, and therefore directly subject to the full spectrum of legal obligations under the GDPR. In the ML context, the business owner/customer typically is the data controller and the IT vendor is the data processor if it processes the personal information solely to provide the service to its customer. To the extent the IT vendor is permitted to process the personal information (e.g., the training data) for its own purposes (such as to improve its service and offer an enhanced product to other customers), the IT vendor may become a data controller with respect to such information, and must comply with the relevant rights.

In light of the different obligations and restrictions that may apply depending on the role of a party in the ML context, a business owner using an IT vendor to provide ML services should carefully assess its data protection role and that of the IT vendor under applicable data protection law to determine which contractual terms are appropriate.

Under the GDPR, a data processing agreement ("DPA") must be put in place between a data controller and a data processor. The DPA must include the content requirements of Article 28 of the GDPR, which must restrict the processor's use of the personal information, impose on the processor data security requirements, as well as audit and sub-processing requirements, and ultimately require the data processor to return or delete the personal information upon termination of the services, among other issues.

Although it is not legally required to be governed by contract, DPAs often impose an obligation on the data processor to notify personal data breaches to the data controller. As explained above, additional contractual terms may be appropriate in the ML context, such as those addressing the use of human input and the performance of relevant risk assessments and other risk mitigation measures. In addition to the DPA, data transfer considerations will apply where the IT vendor receives or accesses EU customer data from the U.S. or other jurisdiction outside of the EU that has not been recognized by

the European Commission as providing an adequate level of data protection. Following the *Schrems II* case in the EU, a data transfer risk assessment should be conducted on a case-by-case basis to process EU customer data in the United States.[3]

There may be challenges in determining whether the IT vendor offering ML services is actually a controller or a processor (or equivalent under applicable law) when, in addition to providing the agreed-upon service, the IT vendor seeks to derive value from an ML process that uses customer data.

A factor to consider in this assessment is whether the IT vendor seeks to use the personal information received from the customer for the vendor's own purposes, and in particular if the ML model learns from data sets of one customer or by combining data sets from different customers and other available sources. While the IT vendor may derive significant value from owning or having license to use the input and output data of an ML model for its own purposes, in doing so it may be subject to legal obligations imposed on data controllers/businesses under applicable law. The compliance bar is thus heightened, and the need to address data protection and security risks is greater when the IT vendor is not limited to a processor role.

In cases where an IT vendor offering the ML services is a data processor/service provider for certain data processing activities and a data controller/business for other data processing activities, the data protection terms of the contract should clarify the roles of the parties and to whom the relevant compliance burden applies.

CONCLUSION

Training a machine learning model using customer proprietary data raises both IP and data protection considerations. These considerations should be appropriately reflected in the agreements executed between the parties to account for the different elements of the ML process, specify ownership and IP rights, and address applicable data protection considerations.

From an IP perspective, when contracting with an IT vendor that offers ML services, the following key considerations should be assessed:

- Identification of each component of the ML process, including the ML model, training data, methods of training, proprietary business know-how, improvements to the ML model, and output from the ML model;

[3] *See* Court of Justice of the European Union ("CJEU") judgment of July 16, 2020, on *Data Protection Commissioner v. Facebook Ireland Limited, Maximillian Schrems* Case (C-311/18). In its judgment, the CJEU invalidated the EU-U.S. Privacy Shield Framework for the transfer of personal information from the EU to the U.S. and considered that that EU Standard Contractual Clauses ("SCCs") for controller to processor data transfers remain valid, subject to a case-by-case transfer risk assessment and, where necessary, additional safeguards for the transfer.

- Consideration of the business model of each party, including how each party derives revenue and the ability of one party to work with competitors of the other party;
- Consideration of the types of IP rights that protect the components of the ML process, including trade secrets, copyrights, patents and database rights;
- Proper allocation if IP ownership and license rights by contract between the parties for the pre-existing IP and any IP that is developed in the ML process; and
- Coordination of data use rights under IP law and IP contract terms, with data use restrictions imposed by data privacy laws.

From a data protection perspective, agreements with IT vendors in the ML context should address the following issues:

- The data protection role of each party with respect to each relevant ML-related data processing activity, particularly where the IT vendor is permitted to derive value from the processing of customer data and re-use the output data to provide services to other customers;
- The minimum contractual responsibilities of the parties (e.g., controller/processor terms) under applicable law (e.g., CCPA, GDPR) with respect to the relevant ML-related data processing activity;
- Additional contractual terms to address data protection considerations specific to ML in an effort to mitigate risks to personal information in the ML testing, training and deploying phase; and
- Jurisdictional considerations regarding the scope of data protection laws applicable to the agreement, and any applicable international data transfer considerations.