MAYER | BROWN

Fall 2019

Artificial Intelligence & Financial Services

Thought Leadership

Foreword

This booklet collects some of our recent thought leadership at the intersection of artificial intelligence (AI) and financial services. In the pages that follow, Mayer Brown partners provide thoughts on:

- Addressing regulatory, privacy/ cybersecurity, and litigation risks;
- Investing in AI and fintech;
- Advising the board on AI risks and issues; and
- The US federal government's AI strategy.

You will see more from us in this area. The majority of our clients are in financial services, and the financial services sector is focused on AI. According IDC, worldwide spending on AI is predicted to increase 44.0% from 2018 to 2019, with Banking being the second largest user at $5.6 billion. We see every part of the financial services industry being transformed, and we intend to continue to provide thought leadership to help you on that journey.

www.mayerbrown.com

Artificial Intelligence & Financial Services

TABLE OF CONTENTS

AI and Big Data Regulatory Risks Under Banking and Consumer Financial Laws

Melanie Brody Joy Tsai
Eric T. Mitzenmacher

Technological advancements constantly reshape America's banking and consumer finance ecosystem. Today, artificial intelligence ("AI") is among the most intriguing technologies driving financial decision-making. Powerful enough on its own to warrant significant investment, AI has even more transformative potential when coupled with industry momentum toward greater use of "big data" and alternative or non-traditional sources of information.

With material changes in banking processes on the horizon, regulators and industry participants brace themselves for the full impact of AI and big data. This article contributes to ongoing discussion by addressing the increasing regulatory focus on issues unique to, or heightened by, AI and big data. After exploring the rise of regulatory interest in these areas, we address specific regulatory risks under banking and consumer financial laws, regulations, and requirements, including: (i) the Equal Credit Opportunity Act ("ECOA") and fair lending requirements; (ii) the Fair Credit Reporting Act ("FCRA"); (iii) unfair, deceptive, and abusive acts and practices ("UDAAPs"); (iv) information security and consumer privacy; (v) safety and soundness of banking institutions; and (vi) associated vendor management expectations.

Regulators Are Increasingly Interested In AI and Big Data

As the use of AI and big data in financial services gradually becomes an industry norm, regulators have become increasingly interested and also have developed a more sophisticated understanding of the area. Federal and state regulators have now weighed in on various product types and banking processes. While doing so, they have exhibited movement from basic information gathering to a more sophisticated approach to understanding regulatory issues. Regulators have not yet promulgated material regulation specifically addressing AI and big data issues—and such active regulation appears to remain a ways off—but they have arguably moved past infancy in their approaches to such issues.

At the federal level, expressions of regulatory interest have come not only from core banking and consumer financial regulators, but also from calls by the Government Accountability Office ("GAO") for broader interagency coordination on issues related to AI and big data. The Consumer Financial Protection Bureau ("CFPB") has sought industry information on the use of alternative data and modeling techniques in the credit process in a February 2017 Request for Information,[1] and members of the Federal Reserve's Board of Governors ("FRB") have spoken on fair lending and consumer protection risks.[2] These

1 82 Fed. Reg. 1183.

2 Lael Brainard, Member, Federal Reserve Board, Speech at Fintech and the New Financial Landscape: What are we Learning about Artificial Intelligence in Financial Services? (Nov. 13, 2018) *available at*https://www.federalreserve.gov/newsevents/speech/brainard20181113a.htm.

regulators have focused, to date, on questions regarding process transparency, error correction, privacy concerns, and internalized biases, even as they see promise in AI and big data's ability to reduce lending risk and/or open credit markets to previously underserved populations. At the same time, the GAO has issued two reports (in March 2018 and December 2018) promoting or recommending interagency coordination on flexible regulatory standards for nascent financial technology ("Fintech") business models (including through "regulatory sandboxes") and the use of alternative data in underwriting processes.[3]

State regulators have also begun to involve themselves in the national discourse about AI and big data. In doing so, they have staked out similar positions to federal regulators with respect to data gathering and understanding technologies, while remaining skeptical of federal overreach in regulating (or choosing not to regulate) AI-driven processes. Various state Attorneys General, for example, have joined the discussion by opposing revisions to the CFPB's policy on no-action letters due, in part, to concern over the role machine learning could play in replacing certain forms of human interaction in overseeing underwriting questions such as "what data is relevant to a creditworthiness evaluation and how each piece of data should be weighted."[4] In addition, the New York Department of Financial Services ("NYDFS") has moved perhaps as far as any regulator—albeit in the context of life insurance, rather than banking or consumer finance—by issuing two guiding principles on the use of alternative data in life insurance underwriting: (i) that insurers must independently confirm that the data sources do not collect or use prohibited criteria; and (ii) that insurers should be confident that the use of alternative data is demonstrably predictive of mortality risk, and should be able to explain how and why the data is predictive.[5] NYDFS or other regulators may see the next logical step as applying similar requirements to the context of credit underwriting.

Not all regulatory interest is bad news for AI, big data, or the companies staking their economic futures on the two. Despite recognizing certain risks, regulators have also publicly acknowledged empirical evidence indicating potential benefits of AI and big data. The CFPB's Office of Research, for example, predicted that the use of alternative data could expand responsible access to credit to the estimated 45 million consumers who lack traditional credit scores.[6] Supporting that prediction, a white paper published by the Federal Reserve Bank of Philadelphia found statistical evidence that use of nontraditional information from alternative data sources do allow consumers with little or inaccurat credit records, based on FICO scores, to have access to credit;[7] and a study by the Federal Deposit Insurance Corporation ("FDIC") noted that one in five financial institutions cited profitability as a major obstacle to serving underbanked consumers, but that new technologies may enable consumers whose traditional accounts are

3 U.S. Government Accountability Office, GAO-18-254, Financial Technology: Additional Steps by Regulators Could Better Protect Consumers and Aid Regulatory Oversight (Mar. 2018); U.S. Government Accountability Office, GAO-19-111, Financial Technology: Agencies Should Provide Clarification on Lender's Use of Alternative Data (Dec. 2018).

4 New York Office of the Attorney General, Policy on No-Action Letters and the BCFP Product Sandbox (Feb. 11, 2019), https://ag.ny.gov/sites/default/files/cfpb_nal_and_sandbox_comment_final.pdf

5 New York Department of Financial Services Insurance Circular Letter No. 1 (Jan. 18, 2019), https://www.dfs.ny.gov/industry_guidance/circular_letters/cl2019_01

6 Consumer Financial Protection Bureau, Data Point: Credit Invisibles (May 2015), https://files.consumerfinance.gov/f/201505_cfpb_data-point-credit-invisibles.pdf

7 Federal Reserve Bank of Philadelphia, The Roles of Alternative Data and Machine Learning in Fintech Lending (Jan. 2019), https://www.philadelphiafed.org/-/media/research-and-data/publications/working-papers/2018/wp18-15r.pdf

closed for profitability issues to continue to have access to financial services.[8]

Regulators' overall attitude toward AI and big data might best be described as "cautiously optimistic." That positioning, as well as expressions of receptiveness toward further review and research, presents the industry participants with an opportunity to help construct the regulatory landscape that will ultimately govern their use of these technologies and processes. But active participation in the regulatory process requires understanding not only of the technological and business opportunities of AI and big data, but also of the legal requirements regulators are seeking to implement and/or balance.

Regulatory Issues Raised by AI and Big Data Are Diverse and Significant

As previously indicated, AI and big data have transformative potential within the banking and consumer finance industries. They are not merely incremental steps forward for credit practices, but instead are leaps toward new marketing, underwriting, and fraud and risk management approaches. Accordingly, they raise legal and regulatory issues across a variety of banking and consumer financial laws and regulatory expectations. Below, we address particular issues raised in six regulatory areas: (i) ECOA and fair lending; (ii) FCRA; (iii) UDAAPs; (iv) information security and consumer privacy; (v) safety and soundness of banking institutions; and (vi) vendor management.

ECOA and Fair Lending: Can Biases Be Controlled and Outcomes Explained?

As financial institutions increase their use of AI in marketing, underwriting, and account management activities, decision-making that is removed from—or at least less comprehensively controlled by—human interaction raises the risk of discrimination in fact patterns that courts and regulators have not previously addressed. Use of big data inputs for credit-related decision-making raises further the risk that new data points, not facially discriminatory, may be relied on by AI as proxies for protected class status.

With respect to federal consumer financial laws, ECOA prohibits a person from discriminating against an applicant on a prohibited basis regarding any aspect of a credit transaction or from making statements that would discourage on a prohibited basis a reasonable person from making or pursuing a credit application.[9] There are two theories of liability under ECOA: (i) disparate treatment, where a creditor treats an applicant differently based on a prohibited basis; and (ii) disparate impact, where a creditor uses a facially neutral policy or practice that has an adverse impact on a prohibited basis, unless the policy or practice serves a legitimate business need that cannot reasonably be achieved by another less discriminatory means. For mortgage loans, the Fair Housing Act imposes similar anti-discrimination requirements, albeit in connection with somewhat different prohibited bases.

States may also impose fair lending requirements, or even fair commerce requirements, that extend beyond lending activities. While such laws frequently protect similar classes as federal fair lending requirements do, some states add protected classes

8 Federal Deposit Insurance Corp., Assessing the Economic Inclusion of Potential of Mobile Financial Services (June 30, 2014), https://www.fdic.gov/consumers/community/mobile/mobile-financial-services.pdf

9 12 C.F.R. § 1002.4.

such as military servicemembers, or expressly protect consumers on the basis of sexual orientation in a manner that may only be implied by federal fair lending requirements.

Regulators have seized on the power of AI to detect patterns in data that may result in unlawful discrimination where traditional underwriting regimes may either have controlled more thoroughly for fair lending risk or simply not identified a pattern on which to make credit-related decisions in the first place. At a November 2018 Fintech conference on the benefits of AI, for example, Lael Brainard, a member of the FRB, noted that firms view artificial intelligence as having superior pattern recognition ability, potential cost efficiencies, greater accuracy in processing, better predictive power, and improved capacity to accommodate large and unstructured data sets,[10] but cautioned that AI presents fair lending and consumer protection risks because "algorithms and models reflect the goals and perspectives of those who develop them as well as the data that trains them and, as a result, artificial intelligence tools can reflect or 'learn' the biases of the society in which they were created." Brainard cited the example of an AI hiring tool trained with a data set of resumes of past successful hires that subsequently developed a bias against female applicants because the data set that was used predominantly consisted of resumes from male applicants. In a white paper, "Opportunities and Challenges in Online Marketplace Lending," the Treasury Department recognized this same risk, noting that data-driven algorithms present potential risk of disparate impact in credit outcomes and fair lending violations, particularly as applicants do not have the opportunity to check and correct data points used in the credit assessment process.[11]

State regulators have also focused on discrimination risk when AI and/or big data are used in underwriting or similar practices. Attorneys General of several states in an October 2018 letter to the Federal Trade Commission ("FTC") commented that the use of AI tools may lead to price-discrimination or price-targeting with negative distributional consequences for certain protected classes of consumers.[12] In addition, while in a different commercial context, the NYDFS recently issued guidance on the use of alternative data in underwriting insurance.[13] Following an investigation into insurance underwriting guidelines and practices, NYDFS identified the same concerns that federal regulators raised—the potential for violations of anti-discrimination law and the lack of transparency for consumers.

The use of AI and big data may present fair lending concerns at all phases of a credit transaction. Federal Reserve staff commented that at the credit marketing phase, the use of big data to determine what content consumers are shown may present redlining and steering risks.[14] An Internet user's web browsing history affects the advertisements he or she is shown as some companies use algorithms to send targeted advertisements. Similarly, companies could use big data to target certain groups of consumers for particular credit products. At the credit underwriting

10 Lael Brainard, Member, Federal Reserve Board, Speech at Fintech and the New Financial Landscape: What are We Learning about Artificial Intelligence In Financial Services? (Nov. 13, 2018) *available at* https://www.federalreserve.gov/newsevents/speech/brainard20181113a.htm.

11 U.S. Department of Treasury, Opportunities and Challenges in Online Marketplace Lending (May 10, 2016), https://www.treasury.gov/connect/blog/documents/opportunities_and_challenges_in_online_marketplace_lending_white_paper.pdf

12 New York Office of the Attorney General, Comment Letter on Competition and Consumer Protection in the 21st Century (Oct. 10, 2018), https://oag.ca.gov/system/files/attachments/press-docs/10.10.2018-multistate-ag-letter-ftc-re-hearings.pdf

13 New York Department of Financial Services Insurance Circular Letter No. 1 (Jan. 18, 2019), https://www.dfs.ny.gov/industry_guidance/circular_letters/cl2019_01.

14 Carol A. Evans, *Keeping Fintech Fair: Thinking about Fair Lending and UDAP Risks*, Consumer Compliance Outlook (2017).

phase, AI models may use alternative data to determine whether to grant credit or to make pricing decisions. Some data points, such as a consumer's educational background or spending habits, may have a nexus with creditworthiness but may also be correlated with race or other prohibited bases. AI algorithms could also use alternative data at the credit servicing phase to determine what modifications to offer a financially distressed consumer or when to engage in account management activities.

Regulators may expect financial institutions that use AI to implement monitoring programs to determine whether their credit models may lead to disproportionate negative effects on protected classes. The CFPB has granted a no-action letter to a company that considers educational information, in addition to traditional credit factors, in underwriting and pricing loans but has also conditioned the no-action letter with commitments to a confidential compliance plan.[15] In surveying companies that use alternative data in credit underwriting, the GAO noted that one Fintech lender monitors the effects any changes to its underwriting models may have on fair lending risk. Some of the lenders surveyed tested their credit models for accuracy, and all discussed testing to control for fair lending risk."[16]

Even in the absence of discriminatory intent or outcomes, AI may complicate compliance with technical aspects of federal and state fair lending requirements. Black box AI systems may make it difficult or impossible for certain financial institutions to comply with adverse action notice or recordkeeping requirements, for example.

With respect to required notifications, ECOA and Regulation B require that creditors provide certain notices regarding actions taken on applications for credit. Adverse action notices must contain either a statement of specific reasons for the action taken or a disclosure of the applicant's right to a statement of specific reasons taken within 30 days if the statement is requested within 60 days of the creditor's notification.[17] Whether provided upfront or only upon consumer request, a creditor's list of reasons for adverse action "must be specific and indicate the principal reason(s) for the adverse action. Statements that the adverse action was based on the creditor's internal standards or policies or that the applicant…failed to achieve a qualifying score on the creditor's credit scoring system are insufficient."[18] The regulatory language would suggest that a generic explanation such as "our proprietary algorithm for credit underwriting determined that you are ineligible" would be insufficient. In contrast, a notice indicating "your credit score is too low," but coupled with reasons for the credit score would likely be deemed sufficiently specific. The Interpretative Guidance to Regulation B further provides that specific reasons disclosed "must relate to and accurately describe the factors actually considered or scored by a creditor." If the creditor bases the adverse action on a credit scoring system, the reasons disclosed must relate only to those factors actually scored in the system. Moreover, no factor that was a principal reason for denial may be excluded from disclosure even if the relationship of that factor to predicting creditworthiness may not be clear to the applicant. Financial institutions using less transparent AI systems may find it difficult to populate an appropriate list of reasons for adverse action and

15 Consumer Financial Protection Bureau, No-Action Letter to Upstart (Sept. 14, 2017), https://files.consumerfinance.gov/f/documents/201709_cfpb_upstart-no-action-letter.pdf

16 U.S. Government Accountability Office, Financial Technology: Additional Steps by Regulators Could Better Protect Consumers and Aid Regulatory Oversight (Mar. 2018), https://www.gao.gov/assets/700/690803.pdf

17 12 C.F.R. § 1002.9(a)(2).

18 *Id.* § 1002.9(b)(2).

those with more transparent AI systems may find themselves responding to consumer inquiries or complaints about credit decisions made on seemingly irrelevant data points over which an AI happened to find a correlation with default rates or other material considerations.[19]

FCRA: When Is "Big Data" a "Consumer Report?"

Big data also presents risks under FCRA, and such risks are amplified if AI-driven underwriting systems have access to alternative data sources without the establishment of proper controls restricting the use of particular data elements. These risks largely relate to financial institutions inadvertently turning information into "consumer reports" under FCRA when neither the financial instruction nor the source of the data intended the data to be subject to FCRA requirements.

FCRA imposes various requirements on persons who provide "consumer reports" (i.e., "consumer reporting agencies"), as well as on persons who use or furnish information for inclusion in "consumer reports." While a traditional consumer credit report is a "consumer report," the term is far broader. Except as expressly exempted, a "consumer report" under FCRA is "the communication of any information by a consumer reporting agency bearing on a consumer's creditworthiness, credit standing, credit capacity, character, general reputation, personal characteristics, or mode of living which is used or expected to be used or collected in whole or in part for determining a consumer's eligibility for credit, employment purposes, or any other purposes enumerated in the statute."[20] (The term "consumer reporting agency" somewhat circularly includes most parties who provide "consumer reports" on a for profit or a cooperative non-provider basis, so the fact that a data source does not consider itself to be a "consumer reporting agency" is not necessarily relevant to a financial institution's obligations when using alternative data.) This broad definition means that a firm that provides data that is actually used for determining credit eligibility may be subject to consumer reporting agency obligations—even if the firm did not intend for the data to be used as such.

Accidentally rendering information from a "consumer report" has a variety of regulatory consequences for a user of alternative data. For example, a consumer reporting agency may furnish (and a person may receive) a consumer report only for "permissible purposes" enumerated under FCRA. For example, a consumer reporting agency may furnish a consumer report to a person who intends to use the report in situations including: (i) in connection with a credit transaction involving the consumer, (ii) for employment purposes, (iii) in connection with insurance underwriting, or (iv) in accordance with the consumer's written instructions.[21] In many cases, entities that obtain alternative data may not have a permissible purpose. In addition, FCRA imposes an adverse action notice requirement (similar to the notice requirements under ECOA) for entities that take action with respect to any consumer that is based in whole or in part on any information contained in a consumer report.[22] Entities that use AI algorithms for credit decisions may have difficulty providing information required in FCRA adverse action notices (such as the specific source of the consumer report and the factors affecting any credit scoring model used in underwriting credit) when it is

19 FCRA also requires users of consumer reports to issue adverse action notices that include specific disclosures regarding numeric credit scores when such scores are used in deciding to take adverse action. 15 U.S.C. § 1681m.

20 *Id.* § 1681a(d)(1).

21 *Id.* § 1681b(a)(3).

22 *Id.* § 1681b(b)(3).

unclear what data points comprise of the consumer report.

Inadvertently converting a data source into a consumer reporting agency also has significant repercussions for the data source. A consumer reporting agency is subject to specific legal obligations, such as obtaining certain certifications from users of consumer reports, ensuring the accuracy of consumer information, investigating consumer disputes of inaccurate information, and filtering out certain items that cannot be reported. The GAO recognized that Fintech lenders who use alternative data in credit underwriting may have sensitive data, such as consumers' educational background or utility payment information, that may contain errors and cannot be disputed.[23]

To protect itself from becoming a consumer reporting agency (and subject to FCRA's numerous obligations), some data sources may include in their service agreements a representation that the firm will not use data for credit underwriting. If the user relies on AI models that, unknown to (or uncontrolled by) the user, pull data points from such a data source, the service agreement representation might be false. If the data used reflects on FCRA-regulated characteristics (e.g., the consumer's creditworthiness, credit standing, reputation, etc.) such that its use in credit underwriting renders the information a "consumer report," the false representation to the data source may be a false certification to a consumer reporting agency for the purpose of obtaining a consumer report. In that circumstance, in addition to possible remedies for breach of contract and regulatory action against the user, FCRA provides the consumer reporting agency a private right of action for such false representations if the representations are willful. Liability under that right of action is the greater of $1,000 or the actual damages suffered by the consumer reporting agency.[24]

Unfair or Deceptive Acts or Practices: Are AI Decisions Consistent with Disclosures?

In addition to potential ECOA and FCRA risk, an entity's use of AI and machine learning may also present risk under the catch-all prohibition against UDAAPs or, in contexts not governed by CFPB's UDAAP standards, the FTC's unfair and deceptive acts and practices ("UDAP") authority. For example, the FTC and FDIC have pursued an enforcement action against a provider of credit cards to consumers with poor credit histories for alleged violations, including a UDAP prohibition for failing to disclose to consumers that certain purchases that triggered the company's risk algorithm could reduce the consumer's credit limit.[25] The company used a behavioral scoring model that penalized consumers for using the credit card for transactions with certain merchants such as marriage counselors, automobile tire retreading and repair shops, and pawn shops. The complaint did not discuss whether certain transactions were reliably correlated with creditworthiness, but appeared more concerned with the fact that use of the behavioral scoring model was not disclosed. As black box AI systems become more prevalent, and such systems may train themselves to use novel algorithms and approaches to underwriting and account management, financial institutions may want to consider the need for broader disclaimers regarding the factors that may impact credit decisions and/or the processes that may develop new approaches to creditworthiness analysis altogether.

23 U.S. Government Accountability Office, GAO-19-111, Financial Technology: Agencies Should Provide Clarification on Lender's Use of Alternative Data (Dec. 2018).

24 15 U.S.C. § 1681n(b).

25 *Fed. Trade Comm'n v. CompuCredit Corp.*, No. 1:08-CV-1976-BBM-RGV (N.D. Ga. 2008), https://www.ftc.gov/sites/default/files/documents/cases/2008/06/080610compucreditcmplt.pdf

Information Security and Consumer Privacy: When Is Big Data Too Big?

Regulators are also aware of heightened cybersecurity and information privacy risks involved with the use of big data (whether in connection with AI-driven processes or otherwise). A GAO report explained that Fintech firms may pose consumer privacy concerns because they collect more consumer data than traditional firms. For example, firms that use alternate data in credit underwriting may have non-public personal information about consumers' educational background, bill payment history, or other sensitive data.[26] The multi-state Attorneys General in a letter to the FTC expressed concern that some firms may be accumulating big data against consumers' wishes "on account of a lack of choice and immense imbalances in market power between service providers and consumers. Consumers often concede valuable competitive data and their privacy interests because they in practice have no choice, other than foregoing the service altogether."[27] A data breach could expose sensitive personal information that consumers did not even want to share in the first place.[28] Financial institutions information security and consumer privacy practices should consider the risks raised by reliance on big data, as well as the extent to which AI-driven processes are able to seek out and utilize/store new forms of data that the financial institution otherwise does not collect.

Safety and Soundness: Can You Demonstrate Your Approach Controls Risk?

When AI and big data processes are used by banking entities, regulators have rounded out their concern about the direct effects of such processes on risk with references to general safety and soundness standards. In a Supervision and Regulation Letter, the FRB emphasized the need for critical analysis through the development, implementation, and use of models for safety and soundness.[29] A GAO report noted that the use of alternative data in underwriting decisions has not been tested in an economic downturn.[30] Some of these concerns may lessen over time, as AI approaches gain a greater history across different timeframes and fact patterns. (While some back-testing may be possible to alleviate regulators' concerns, the historic availability of alternative data with which to conduct tests across different macroeconomic climates—for example—may not be as robust as the historic availability of traditional credit data.) Until that point, however, regulators seem to expect AI risk to be monitored and controlled similarly to traditional credit practices.

Vendor Management: Can You Understand and Control Vendors' AI and Big Data Use?

Finally, beyond direct concerns as to violations of law and control of risk by financial institutions themselves, regulators have expressed interest in limiting the risk that financial institutions expose

---

26 U.S. Government Accountability Office, GAO-18-254, Financial Technology: Additional Steps by Regulators Could Better Protect Consumers and Aid Regulatory Oversight (Mar. 2018).

27 New York Office of the Attorney General, Comment Letter on Competition and Consumer Protection in the 21st Century (Oct. 10, 2018), https://oag.ca.gov/system/files/attachments/press-docs/10.10.2018-multistate-ag-letter-ftc-re-hearings.pdf.

28 On the other hand, the FRB has implicitly acknowledged the power of AI in fighting cyberattacks by suggesting that supervised institutions may need to develop their own AI tools to identify and combat outside AI-powered threats. Lael Brainard, Member, Federal Reserve Board, Speech at Fintech and the New Financial Landscape: What are we Learning about Artificial Intelligence in Financial Services? (Nov. 13, 2018) *available at* https://www.federalreserve.gov/newsevents/speech/brainard20181113a.htm.

29 Federal Reserve Board, SR Letter 11-7, Guidance on Model Risk Management (Apr. 4, 2011), https://www.federalreserve.gov/supervisionreg/srletters/sr1107.pdf

30 U.S. Government Accountability Office, GAO-19-111, Financial Technology: Agencies Should Provide Clarification on Lender's Use of Alternative Data (Dec. 2018).

themselves and/or consumers through partnerships with vendors who may rely on AI or big data processes. The FDIC,[31] OCC,[32] FRB,[33] and other supervisory regulators have long-expected financial institutions to control for risks involved in third-party vendor relationships and have issued guidance on effective third-party risk management. Management of vendors use of AI and big data is merely another prong in effective vendor oversight. That said, vendors may consider their systems proprietary and confidential or may otherwise maintain "black box" AI systems that cannot be fully explained. The FRB acknowledged that "it is not uncommon for there to be questions as to what level of understanding a bank should have of its vendors' models, due to the balancing of risk management, on the one hand, and protection of proprietary information, on the other. To some degree, the opacity of AI products can be seen as an extension of this balancing, but AI can introduce additional complexity because many AI tools and .models develop analysis, arrive at conclusions, or recommend decisions that may be hard to explain to regulators."[34] More concretely, NYDFS has taken the position that an insurer "may not rely on the proprietary nature of a third-party vendor's algorithmic process to justify the lack of specificity related to an adverse underwriting action,"[35] and that expectation to understand a vendor's AI models could also apply to the context of credit underwriting.

Most regulatory guidance on third-party risk management does not specifically address the challenges of understanding AI. For example, the FDIC guidance discusses risks that may be associated with third-party lending arrangements, as well as its expectation that financial institutions implement a process for evaluating and monitoring vendor relationships that include risk assessment, due diligence, contract structuring and review, and oversight.[36] However, the OCC has issued an FAQ that specifies that relationships between Fintech companies and banks may be subject to its bulletin on vendor risk management.[37] The OCC acknowledged that a bank may not be able to receive in-depth information on every third-party service provider that supports critical activities, but the OCC nonetheless expects the bank to: (i) develop appropriate alternative ways to analyze critical third-party service providers; (ii) establish risk-mitigating controls; (iii) be prepared to address interruptions in delivery; (iv) make risk-based decisions that the critical third-party vendors are the best service providers available despite the bank's inability to acquire all the information it seeks; and (v) retain appropriate documentation of efforts to obtain information.[38]

---

31 Federal Deposit Insurance Corporation, Examination Guidance for Third-Party Lending (July 29, 2016), https://www.fdic.gov/news/news/financial/2016/fil16050a.pdf

32 Office of the Comptroller of the Currency, Risk Management Guidance, 2013-29 (Oct. 30, 2013), https://www.occ.gov/news-issuances/bulletins/2013/bulletin-2013-29.html

33 Federal Reserve Board, Guidance on Managing Outsourcing Risk (Dec. 5, 2013), https://www.federalreserve.gov/supervisionreg/srletters/sr1319a1.pdf

34 Lael Brainard, Member, Federal Reserve Board, Speech at Fintech and the New Financial Landscape: What are we Learning about Artificial Intelligence in Financial Services? (Nov. 13, 2018) *available at* https://www.federalreserve.gov/newsevents/speech/brainard20181113a.htm.

35 New York Department of Financial Services Insurance Circular Letter No. 1 (Jan. 18, 2019), https://www.dfs.ny.gov/industry_guidance/circular_letters/cl2019_01.

36 Federal Deposit Insurance Corporation, Examination Guidance for Third-Party Lending (July 29, 2016), https://www.fdic.gov/news/news/financial/2016/fil16050a.pdf.

37 Office of the Comptroller of the Currency, Frequently Asked Questions to Supplement OCC Bulletin 2013-29 (June 7, 2017), https://www.occ.gov/news-issuances/bulletins/2017/bulletin-2017-21.html.

38 *Id.*

Conclusion

While advances in technology show a lot of promise for the financial services industry, many regulators have raised questions about responsible use from the consumer protection perspective. Regulators have developed an improved understanding of AI and machine learning, but they are also receptive to gathering more information to develop standards governing the industry. The banking and consumer finance industries are at a crucial point in the development of AI and big data processes. Careful engagement with regulatory issues raised by new technology and practices across a range of requirements and contexts will be important to the development and expansion of sustainable credit programs built around significant reliance on AI and big data.

Explainable AI (XAI) and Litigation Defensibilty

Eric B. Evans
Alex C. Lakatos
Brad L. Peterson

Machine learning algorithms and other applications of artificial intelligence are making more and more day-to-day business decisions. Thirty years ago, if an entrepreneur wanted a loan for her startup, she'd walk into a bank and talk to a loan officer. Ten or twenty years ago, she might apply online, but a loan officer would still have final approval. Now, a machine learning algorithm will often make the call. Same for job applicants: thirty years ago, they'd apply by mail or in person. Ten or twenty years ago, online, to a human decision maker. And now, their applications feed into machine learning systems that make the key calls.

Our legal system is evolving. It has elaborate rules governing how to prove what people did and when. It has long-established assumptions about who legal actors are and how to find their intent. But these rules assume that human are the last step in a decision process. When a machine learning system or another form of artificial intelligence is the final step in the decision process, these assumptions break down.

At a basic level, if our entrepreneur doesn't get a loan today, she can't ask the machine why. Same for the job applicant: the machine won't have an answer. This is true even if the bank or employer keeps a human in the loop at the end. That person will only be able to say *that* the machine made a determination.

And just as that answer isn't likely to satisfy the entrepreneur or the job applicant, it isn't likely to satisfy a judge, jury or litigation opponent either.

The need to—and the difficulty of—explaining how business-critical artificial intelligence systems work is a key new challenge for companies that rely on them.

"Artificial intelligence" brings to mind sentient robots, such as the heroic Mr. Data of *Star Trek* and numerous robotic villains. But today, in reality, it means software systems applying complex mathematics to predict outcomes based on data fed into them. Generally, AI users will want to prove that the business decisions implement policy choices made by company management and the choice of algorithms and parameters by data scientists and programmers based on those policy choices. The AI is not a decision maker but merely a mechanism for implementing business decisions.

Businesses now recognize that business-critical tools that only highly trained experts can explain create regulatory and litigation risk. Regulators aren't likely to be satisfied by pointing to a machine learning tool than to explain a rejected application. And most litigation revolves around business decisions. The parties will take and defend depositions of the key human decision makers. They will collect, review and produce the documents these decision makers created. These procedures are well-defined. But there's no way to depose an artificial intelligence tool, so a company that allows its tool to be seen as the decision maker will have difficulty defending the decisions. And the inputs and outputs that reflect its operation will not be decipherable by judges, juries or non-experts.

As a result, many companies are focusing on ***explainability.*** Explainability, in general terms, has three aspects:

- **Transparency:** easy identification of the important factors in the tool's operation;

- **Interpretability:** easy identification and explanation of how the tool weights those factors and derives them from its input data; and
- **Provenance:** easy identification of where input data originated.

When an AI tool has all three aspects, a company can explain its results to a regulator, judge or jury in plain language. That is, it can say, "The tool came to this result because it took these inputs, applied these weights to them and derived this result." To achieve that, we recommend that:

- The management team clearly specifies to its data scientists and technicians how the company wants to tool to work, recognizing that those specifications are, in fact, the business decisions;
- The tool is built to store the right facts about how it arrived at results in manner approved by your e-discovery/information governance (EDIG) team;
- The company employs "AI sustainers" to continually test and modify the tool to keep it working as the management team intended; and
- The company employs "AI explainers," people who know how to explain the tool's results.

In a litigation, then, explainability fades into ***defensibility.*** The "AI explainers" within the company will be able to use data retained by the EDIG group to explain how the decisions reflect corporate policy. For plaintiffs, it may be very difficult to find an expert who can speak with authority on an extremely complex proprietary AI tool, at least compared to a data scientist, an AI explainer or an AI sustainer who was part of the team building and sustaining the tool. From a litigation perspective, explainability allows humans to take the witness stand to defend business decisions implemented through an AI tool instead of having the plaintiff's counsel claim that the company is responsible for how a villainous robot abused the plaintiff. The issues shift to the more defensible questions of whether management chose the right policies and the technicians configured the tool correctly.

Investing In AI Fintech Companies

Amanda Baker
Joe M. Pennell
Rebecca S. Eisner
Elizabeth A. Raymond*

Transactions involving fintech companies, and particularly fintech companies incorporating artificial intelligence ("AI") into their products and services, are now commonplace in the fintech landscape. *CB Insights* reports that AI startups are emerging at record rates, with 1,800 new startups raising equity for the first time since 2016, $19 billion of equity funding in 2018 and more than 5,000 equity deals across multiple industries since 2013[39]. Legal and business transaction leaders should carefully consider the range of possible investments in companies offering AI products and services relating to financial services ("AI fintech companies"), and the potential risk and rewards of these investments.

For purposes of this article, we reference a Deloitte definition of artificial intelligence as "the theory and development of computer systems able to perform tasks that normally require human intelligence."[40] AI has the potential (or likelihood) to transform the provision of financial services. Large financial institutions have traditionally been hampered by their legacy technology systems and cumbersome physical operations as well as the need to comply with complex and evolving regulatory requirements. As a result, a consistent theme is that incumbent financial institutions will need to collaborate with their AI fintech company disrupters, using commercial arrangements, partnerships and acquisitions to remain competitive. Incumbent financial institutions have advantages of their own, including large financial resources, the massive ability to manufacture compliant financial products, a wealth of data about their customers' financial activities and the deep-seated trust of their customer base, including tech-savvy millennials. Given the increasing speed at which AI and fintech are developing, the older sourcing strategies of "build versus buy" are being replaced with strategies that allow for flexible and rapid collaboration across a variety of licensing and capabilities acquisition models.

In this article, we will review the spectrum of possible AI investments -- ranging from licensing and service agreements to platform collaborations to financing transactions to joint ventures and strategic partnerships to minority and majority investments and, finally, to M&A-style acquisitions. We will also outline some of the due diligence, structure and contractual considerations for each type of transaction. We will focus on these considerations from the point of view of the buyer of, the investor in, the customer of or the lender to an AI fintech company, with potential AI fintech company counterparties including AI software licensors, cloud-based AI providers, financial data and analytics companies, and AI fintech platform companies. As described in this article, along this spectrum the financial institution may license AI technology, enter into an AI technology services agreement, enter into a "powered by" or white label commercial agreement, provide financing to or purchase whole loan assets of

* Mses Baker, Eisner and Raymond and Mr. Pennell are partners at Mayer Brown LLP. The authors gratefully acknowledge the assistance of Corina Cercelaru, Lawrence R. Hamilton, Joshua La Vigne and Donald S. Waack in preparing this article.

39 CB Insights, *What's next in AI?*, www.cbinsights.com, page 10.

40 https://www2.deloitte.com/content/dam/Deloitte/in/Documents/strategy/in-strategy-innovation-artificial-intelligence-next-bold-play-noexp.pdf page 4.

the AI fintech company, purchase a minority stake in or joint venture with the AI fintech company or acquire a majority interest in or all of the AI fintech company in an M&A transaction.

AI Licenses and Service Agreements

Licensing AI capability from an AI fintech company through a license or service arrangement is likely the fastest way to obtain AI for use by financial institutions. This may take the form of an on-premises license of AI that will be installed, trained and operated by the financial institution, or it may be offered as a "Software as a Service" solution in the cloud by the provider.

Many financial institutions are turning to a collection of AI fintech providers to test the waters. A good, lower risk way to do this is through a "proof of concept" arrangement. A proof of concept arrangement is a short-term agreement that allows a financial institution to test, and an AI fintech company supplier to prove, the value of an AI product or service.

Once the proof of concept is complete, the financial institution may license the AI from an AI fintech provider. Financial institutions should seek to satisfy the usual requirements for critical third-party service provider agreements in their AI licenses and services agreements. AI licenses present a few unique topics, including legal compliance of the AI decisions, allocating ownership and use rights of the components of AI, data use and privacy, and protection of intellectual property rights.

Legal Compliance. First and foremost, AI-based decisions must satisfy the laws and regulations that apply to financial services. This requires the financial institution to apply the same level of diligence to the AI tool or service that the financial institution applies to its other critical third-party products and services.

Of particular concern is that AI-based decisions may discriminate because they rely on data that reflects a discriminatory past or looks only at correlation instead of causal factors. Financial institutions that use AI tools in credit decisions or fraud detection, for example, must ensure that these tools do not discriminate against certain protected classes of applicants or employees. AI tools used for insurance decisions will have to follow recently issued requirements from the New York Department of Financial Services on the use of "unconventional sources or types of external data" to address the risk of unlawful discrimination and a lack of data transparency.

In addition, AI systems should produce output that is transparent, auditable and that can be explained – sometimes called "Explainable AI." Licenses from an AI fintech company should address the extent to which the AI decisions and outcomes are explainable, and the method by which the financial institution may access those explanations and related data. The license agreement may also need to specify that the AI may be subject to regulatory examination, and require the AI fintech provider to cooperate with such examinations. Financial institutions may also want to require that AI has "circuit breakers" – a method for pausing operations to gather data about correct and compliant operation, confirm security compliance, and make necessary adjustments in the AI tool to eliminate errors, mistakes and bias. Record-keeping and audit requirements are also important considerations for financial institutions. Because AI tools evolve, data sets change and iterations are part of the process, financial institutions should address how they can access versions of past decisions based on AI tools and data sets that have shifted over time. This is particularly important when financial institutions are using AI in a provider cloud and when the financial institution is not in control of archiving the AI components and outputs.

Financial institutions can mitigate these AI risks by utilizing oversight, risk management and controls to meet legal compliance and business objectives, and by incorporating provisions addressing these requirements in the AI license. Finally, consider whether financial institutions should include rights to training and access to specialists who are familiar with the AI tools and can assist the financial institution with its training, use and ongoing monitoring requirements. Regular compliance meetings with the provider may be required to provide assurance on these key items.

Allocating Ownership and Use Rights and Training Obligations. There are a number of important questions for financial institutions to consider regarding contractual ownership and use of the components of AI in their licensing agreements. These components include the AI tool, evolutionary changes to the AI tool, the training data and instructions, and the output of operation of the AI tool. When licensing AI, most AI fintech providers will expect to continue to own the underlying AI tool, and some may expect to own the evolutionary changes as well. Much of the AI that financial institutions will use may require training. The license should address which party will train the AI, which party will own the training instructions, and which party will own the evolutionary works of the AI tool based on the training. Shifting to the output of the AI tool, most financial institutions would expect to own the decisions and the decision criteria of the AI tool, and this must be specified in the license agreement to achieve that result. Once the parties have determined how they will allocate these ownership rights, they also need to determine whether and to what extent the other party will have ongoing license and use rights in those components.

Data Use and Privacy. Data is the fuel for AI, but data use must comply with the privacy, data security, export control and other laws that apply to the data. In addition, data use must comply with any contractual requirements to third-party data suppliers. These are often not well understood. To guard against these data pitfalls, financial institutions should inquire as to the level of legal and regulatory diligence that has been done on the uses of data to fuel AI systems. The license should specify whether the AI will rely on provider data or financial institution data or both, and, importantly, which party will own which data, and which party may use that data and for what purposes. The license agreement may also specify that the party supplying the data is responsible for obtaining necessary consents and rights to use that data for the AI, and address liability for issues arising from improper use or failure to obtain proper consents. If financial institution data are used for the AI, and those data include non-public personal data, the financial institution will have to assess compliance with its privacy policies governing that data. Similarly, many countries, such as European countries, have tough data protection laws that prohibit the use of individual data for automated processing to evaluate any feature of behavior, preferences or location absent the explicit consent of the individual, and yet, automated processing of individual data to determine preferences is the hallmark of many AI tools. Consider whether the license should require the provider to conduct privacy assessments of the AI tool on a periodic basis.

Protection of IP Rights. Patent, copyright, trade secret and other IP laws were written with a bias to protecting human creativity. Intellectual property (IP) laws in the United States do not square nicely with AI. Not only may a financial institution not own AI that it pays to create, it also may not have the means to fully protect its AI under U.S. IP laws. Contractual protections are a key element of capturing and preserving value in the creation of and returns on the investment in AI. These protections, to be effective, must be implemented before the AI effort begins, and will rely on clear statements of ownership and use

rights in the various components of AI as addressed above.

Service agreements in which AI fintech providers use or rely on AI on an incidental basis to deliver the services are another channel through which financial institutions may obtain the use of AI. Although the AI may not be the cornerstone of such an arrangement, financial institutions should require service providers to reveal if they are using AI tools to provide the services, and if so, they should understand the uses. If the uses bear on any of the issues described in the prior paragraphs above, then the financial institution should take care to perform diligence on those uses, and to define the contractual rights and obligations with respect to such AI as part of the service agreement.

Platform Collaboration and White Label Arrangements

Broadly defined, a digital platform is an integrated framework of digital tools and services that implements key business processes to facilitate exchanges between producers/manufacturers of services/products and consumers. Put more simply, it is the foundation on which a digital business is built. The difference between digital platforms and previous methods of technological transformation is the ***exchange*** function of a platform. In addition, platforms are not simply cost-saving technology for companies' back-office functions; instead, digital platform technology is transforming customer-facing, revenue-generating functions.

These exchanges vary in terms of openness and complexity. While one often thinks of platforms as vast "many-to-many" systems (e.g., Facebook, Google, AirBnB, Uber), a platform can also include more traditional exchanges where a single producer is trying to reach many consumers. These traditional exchanges are exemplified by fintech platforms, which can include systems for consumer banking (e.g, SoFi, Stash), retail investment (Robinhood), payments (Venmo, PayPal, Zelle), loan origination (Lending Tree and multiple white label lending platforms for individual banks), and financial advice (Robo-advisors, H&R Block, Watson).

A financial institution may seek to enter into a commercial arrangement to white label an AI fintech company's digital lending or other digital financial services platform for use by the financial institution. These commercial arrangements – known as a "platform collaboration" or a "white label" or "powered by" arrangement – allow the financial institution to obtain AI capabilities as opposed to building its own AI capabilities. Some of the advantages of a platform collaboration include: (1) relatively small investment for the financial institution; (2) the financial institution gains efficiency because it is not "reinventing the wheel" where AI solutions may already exist in other formats; (3) faster time to enter the market because developing AI is outside of the financial institution's core competency; (4) the financial institution can focus on its core competency; and (5) platform or white label arrangements allow for scalability.

There are, however, risks and disadvantages that must be addressed in any platform collaboration. As noted in the "AI Licenses and Service Agreements" section above, data security and privacy are major issues that the financial institution must consider carefully. Further, under this type of arrangement, the financial institution may have very little control over the direction of the AI platform. Lastly, AI fintech company providers are often time-hungry, highly leveraged start-ups seeking to maximize the rapid growth that follows from successful early entry into an AI fintech company space. Thus, the financial institution must consider the financial stability of the AI fintech company provider and include adequate protections in the contract (e.g., termination rights for

financial degradation, rights to retrieve data in a usable format upon request and termination assistance rights to facilitate a smooth transition to an alternative platform).

In negotiating a platform collaboration contract or a white label arrangement for AI capabilities, a financial institution may find it helpful to leverage its experience from contracting with providers of ERP, information technology (IT) infrastructure and other back-end technology services. In fact, many of the concerns described above in the "AI Licenses and Service Agreements" section are also present in platform collaboration and white label arrangements.

Even the most experienced financial institutions, however, will face unique issues when it comes to platform collaboration deals or white label arrangements for AI capabilities. One such issue is legal compliance. Similar to licensing agreements for AI capabilities, the financial institution must ensure that the white label services and the financial services platform (including the AI tool) comply with all laws and regulations that apply to financial services. The AI fintech provider will most likely try to limit its obligations to complying with laws applicable to the AI fintech provider in its provision of the services. That universe of laws is generally small, and the financial institution may seek to allocate more responsibility on the AI fintech provider for violations of laws applicable to the financial institution that are caused by the AI fintech provider. The parties will need to find a middle-of-the-road approach that provides adequate protection for the financial institution. One compromise for the financial institution to consider is to require the AI fintech provider to bear responsibility for (a) complying with laws applicable to the AI fintech provider in its provision of the services and (b) violations of other laws caused by the AI fintech provider's failure to follow the financial institution's written instructions with respect to such other laws. Another compromise is to require the AI fintech provider to bear responsibility for complying with (x) laws applicable to the AI fintech provider in its provision of the services and (y) any laws that are applicable to the financial institution (but not to the AI fintech provider as a technology provider of the services) provided that the financial institution informs the AI fintech provider of such laws in advance.

Another thorny issue in platform collaboration contracts and white label services arrangements is the ownership rights for developed IP. As mentioned in the "AI Licensing and Service Agreements" section, the parties need to clearly allocate IP rights. The parties need to consider who will own the developed IP that incorporates both the financial institution's and the AI fintech provider's proprietary materials. For example, the developed IP may combine fraud models from the AI fintech provider and underwriting criteria and credit policies from the financial institution. If there are practical challenges in separating that combined, developed IP upon termination of the contract, the parties may consider requirements to delete or destroy that IP upon termination. The parties, however, will need to assess this issue on a case-by-case basis, depending on the circumstances of the deal.

Financing AI Fintech Companies

There are a variety of financing options available for financial institutions lending to, or investing in, AI fintech companies. The type of financing that the lender will execute typically relates to the AI fintech company's experience in the finance industry as well as the space in which the AI fintech company wants to brand itself – technology or finance. Assuming that the AI fintech company's business model is to make loans to customers, most AI fintech start-ups and AI fintech companies without extensive experience in the financial services industry enter into whole loan sale transactions with various investors or

lenders before moving onto capital markets transactions. The motivation for the AI fintech company is threefold, as these types of transactions: (1) allow AI fintech companies a flexible relationship with an investor or lender memorialized in a few documents that can easily be amended and do not trigger significant regulatory compliance, (2) provide exposure for the AI fintech company to various investors and lenders and (3) are structured as off-balance sheet for accounting purposes.

From the point of view of both the financial institution acting as the investor or lender and the AI fintech company, whole loan sales with a single investor or lender are not structurally complex transactions that trigger extensive regulatory compliance and diligence. Instead, these transactions are usually structured as a one-time (or multiple, scheduled) sale(s) from the AI fintech company to its investor or lender where the AI fintech company and the investor or lender agree to the sale(s) on certain negotiated terms. Additionally, the AI fintech company agrees to service the assets and undertakes the servicing responsibilities in the transaction documents. Given the nature of AI fintech companies, servicing is a crucial component for the investor or lender to consider in financing transactions. Servicing responsibilities usually include collecting payments from the underlying obligors on the assets, monitoring the activity of the underlying obligors, enforcing the obligor contracts, taking action to maximize collections in the event of obligor delinquency or default, and providing the requested servicing and performance data to the investor or lender. While the AI fintech company does need to comply with its general corporate and licensing regulatory requirements, this structure does not trigger the typical Dodd-Frank regulatory requirements or generally require registration with the Securities and Exchange Commission. Finally, since there are not multiple transaction parties, the AI fintech company and investor or lender can more easily amend the deal documents if changes are needed as the AI fintech company hones its data systems and servicing policies and procedures.

Since whole loan sales can be papered by a handful of documents, an AI fintech company is able to easily enter into multiple transactions with various investors or lenders. By having exposure to various investors and lenders in a whole loan sale program, the AI fintech company accesses liquidity from multiple sources, which also lowers the financing risks for any single investor or lender. Additionally, whole loan sale investors may be non-bank private equity or hedge funds that often seek leverage from larger, more traditional financial institutions, providing exposure for the AI fintech company to financial institutions that the AI fintech company may not be able to obtain on its own. Accordingly, whole loan sales set the stage for more complex financing transactions in the future.

Finally, investors and lenders typically structure whole loan sales as a true sale from the AI fintech company to the third-party investor or lender. This type of transaction is appealing both to the AI fintech company, since it allows it to obtain financing while easily achieving off-balance sheet treatment through a true sale to an unaffiliated third-party investor or lender, and to the investor or lender, since it should provide isolation from bankruptcy risk. By achieving off-balance sheet treatment, AI fintech companies are also more easily able to brand themselves as technology companies rather than companies that operate in the financial services space.

Investors and lenders may also offer their AI fintech companies financing through a warehouse facility. A warehouse facility is typically negotiated between the AI fintech company and an agent bank lender. These types of facilities are often syndicated to a group of investors or lenders through the agent. Additionally, whole loan investors and lenders will often provide financing of the equity piece under these

structures. Warehouse transactions provide investors and lenders with another option to finance AI fintech companies that is slightly more complex than a whole loan sale, but not as sophisticated and regulatory intensive as a capital markets transaction.

While whole loan sales and warehouse loans offer AI fintech companies relatively straightforward access to liquidity from a variety of financing sources without necessitating significant regulatory compliance, it is, nonetheless, advantageous to maximize funding options through a variety of finance transactions. Investors and lenders providing whole loan and warehouse loan facilities will seek the ability to take out their financing through capital markets transactions. While some financial institution investors and lenders may be comfortable purchasing whole loans, others may prefer to purchase securities backed by such loans for risk and liquidity purposes. Thus, in addition to whole loan sales, AI fintech companies may look to access the capital markets and, more specifically, the structured finance markets. While securitization transactions can provide a more efficient cost of funds for the AI fintech company, investors and lenders will require attention to significant additional regulatory requirements and the AI fintech company will need to have adequate legal, compliance, systems and servicing procedures in place to provide the data and access to employees necessary to facilitate compliance. The financial institution acting as investor or lender may also act as underwriter, initial purchaser or placement agent for the securitization. The underwriter will assist the AI fintech company entering into a securitization, which typically requires the following:

- static pool data on prior transactions or vintage data and pool data relevant to the assets included in the transaction;
- customary narrative descriptions of the company's material underwriting and servicing practices, and other written information for use in an offering document, such as disclosure on the legal and business risks relating to AI-based products and services;
- holding 5% of risk in the transaction;
- coordination with accountants to facilitate the provision of a customary agreed upon procedures letter by an independent accounting firm;
- allowing reasonable access for rating agencies and the investment banking firms to the company's origination and servicing personnel and its records relating to the assets to be securitized and employees with responsibility and knowledge with respect to the securitized assets;
- maintenance of a 17g-5 website allowing any nationally recognized rating agency to access information about the transaction;
- undertaking to make certain filings with the Securities and Exchange Commission; and
- additional requirements if the securitization will be a public offering of securities.

As AI fintech companies enter into financing arrangements with third-party investors or lenders, both AI fintech companies and the investors or lenders should consider the different funding options available to a growing AI fintech company. While whole loan sales provide access to liquidity without as many extensive or complex diligence and legal requirements, not all investors and lenders want to hold loans, and warehouse facilities and capital markets transactions typically have a higher dollar amount. Conversely, while warehouse facilities and capital markets transactions require more diligence and regulatory compliance, they offer access to high dollar bond issuances with multiple sophisticated third parties. AI fintech companies without extensive experience in the financial services industry, as well as their investors and lenders, should consider these factors when establishing funding plans.

Joint Ventures and Strategic Partnerships

The term "joint venture" is quite broad and can involve creating a new entity, an ongoing contractual relationship or a combination of both. As distinguished from a strategic investment or an M&A transaction, a joint venture typically involves two or more parties that come together to achieve a common goal for profit.

In the current regulatory environment, it may be relatively rare for large financial institutions to joint venture or partner with an AI fintech company in the traditional sense, but other large non-bank finance companies may consider the joint venture structure attractive. As discussed below, a large financial institution, such as a bank holding company or an insurance company, is typically highly regulated and seeks to avoid obtaining "control" of the AI fintech company, in most cases by keeping a minority equity investment below 5% (or 10% in the case of an insurance company) of the AI fintech company's voting shares and otherwise avoiding indicia of control. Indicia of control include holding a voting seat on the company's board of directors, certain veto or consent rights, entering into a management agreement or entering into significant business or commercial relationships with the AI fintech company. If the financial institution seeks a control relationship, it may be simpler to acquire complete control through an acquisition as opposed to a joint venture or partnership. On the other hand, the financial institution may forego any equity investment in order to avoid these control questions and seek only a commercial or financing arrangement as discussed above.

Assuming that the joint venture partners are willing to have their joint venture entity be treated as a regulated entity or the joint venture entity is otherwise not subject to what may be viewed as burdensome bank or insurance regulations, there can be a number of advantages to using a joint venture entity as opposed to a contractual joint venture. These advantages include: (a) access to technology, subject matter experts like data scientists, and products contributed to the joint venture as well as distribution channels and markets with greater economies of scale; (b) sharing of regulatory risks that accompany financial institutions, especially when entering a new market; (c) internal and external constituencies (e.g., employee talent in the joint venture and end users of the technology) will perceive a separately identifiable and visible enterprise conducting the joint venture business, with the venture lending itself more to AI innovation than to regulated bank or insurance activity; (d) interests in a joint venture are generally easier to sell or transfer than a collection of contractual relationships; (e) the joint venture entity creates an independent vehicle with greater flexibility and convenience for capital-raising activities; (f) the joint venture entity provides a familiar structure (e.g., a corporation, limited liability company or limited partnership) in which management and governance rules can be established and in which directors, officers and employees typically play familiar roles in making decisions and implementing them, with this level of oversight likely being important in the developing area of AI; (g) the joint venture entity provides a convenient vehicle for measuring profits and allocating and distributing them to the joint venture parties; (h) the joint venture entity can be an independent employer providing identification and focus for employees, including incentive compensation such as equity interests and the opportunity to work on cutting-edge AI projects; (i) the joint venture entity largely enables the joint venture parents to insulate themselves from the liabilities of the joint venture business; and (j) the joint venture entity creates the potential for flexibility in addressing tax matters.

Disadvantages of the joint venture structure – other than the perhaps overriding concern that the AI fintech joint venture will become a regulated entity based on its control by a regulated financial institution – include: (a) complexity because establishing a separate joint venture entity often involves initial and ongoing issues, tasks and costs that are not necessarily present in a contractual association, with time-consuming oversight required by senior managers of the alliance participants; (b) a likely more complicated unwind process because assets, contracts, employees and other resources of the joint venture business may be property of, or affiliated with, the joint venture entity; (c) loss of control in that the joint venture business will normally be, in large part, conducted by the joint venture entity and the rights and ability of the joint venture entity and its activities will be limited by the governance rules of the joint venture entity; (d) difficult fiduciary duty and conflict of interest issues may arise with a joint venture entity that may not arise in a contractual joint venture (although these can largely be handled contractually); and (e) the contractual joint venture can allow more flexibility in staging and developing the joint venture by establishing an initial "let's get our feet wet" relationship without the more substantial commitment involved in establishing, and providing assets and other resources to, a separate joint venture.

Stock Investments and M&A Transactions

Strategic investments and M&A transactions offer a large financial institution, such as a bank or insurance company, some additional flexibility to tailor an investment to its specific business strategy, with each structure having its own unique advantages and disadvantages. Two general concerns applicable to each structure are: (1) the "control" analysis described above in the "Joint Ventures and Strategic Partnerships" section and the effect of bank or insurance regulatory control on the AI fintech company; and (2) the level of diligence a potential investor should complete with respect to each structure. In this section, "investor" refers to financial institutions as investors in or acquirers of AI fintech companies.

A passive, non-controlling investment can offer a large financial institution investor and the AI fintech company a number of advantages. These advantages include: (a) allowing the investor to leverage the AI offerings of the AI fintech company in its business with relatively low risk to the investor due to a limited commitment of resources; (b) potentially less stringent due diligence requirements of the AI fintech company, in general, than majority investments and M&A transactions, but this can vary depending on the cost/benefit analysis and risk tolerance of each individual investor; (c) the imposition of fewer regulatory burdens on the AI fintech company; (d) allowing the AI fintech company to leverage the infrastructure and expertise of the investor; and (e) the AI fintech company's retention of a certain level of autonomy. Disadvantages of this structure include: (w) very limited investor control over the AI fintech company's activities (e.g., no board seat, very few consent rights over activities of the AI fintech company, etc.); (x) limited investor protective provisions; (y) requiring the investor to conduct a relatively complex and ongoing control analysis for regulatory purposes; and (z) tension created due to the differing goals of the investor (financial return) and the AI fintech company (long-term viability). The obligation of the investor to continually assess its level of control over the AI fintech company to avoid subjecting the AI fintech company to regulatory oversight is a key disadvantage to a minority investment. For example, a bank holding company investor must ensure its equity investment remains below 5% in addition to monitoring other means of exercising control over the AI fintech company, such

as the appointment of a board member, veto rights over certain actions of the AI fintech company, ownership of 25% or more of any class of voting securities, rights of first refusal and ownership of convertible securities.[41] As a protective measure, a minority bank holding company investor should seek to include certain transfer rights, such as a put right, for itself in connection with its investment to allow the investor to exit the AI fintech company if regulatory concerns arise.

Investments by insurance companies (or their affiliates) will potentially be subject to the laws governing insurance holding company systems in the states where the insurance companies are domiciled (or deemed commercially domiciled). Generally, those laws presume control – and thus an affiliate relationship – to exist where one person, directly or indirectly, owns 10% or more of the voting securities of another person, although that presumption can be rebutted by submitting a disclaimer of control to the domiciliary state insurance commissioner. In addition, other types of rights, such as the appointment of board members, may be deemed by an insurance commissioner to constitute control of an entity. The laws in many states limit the ability of an insurance company to acquire a controlling minority interest in another entity. In addition, if an entity is treated for insurance regulatory purposes as an affiliate of an insurance company, that relationship will need to be disclosed in the insurance company's statutory financial statements, annual holding company registration statements and enterprise risk reports, and the domiciliary state insurance commissioner will need to be notified in advance of material transactions between the insurance company and its affiliate, giving the commissioner an opportunity to review the transaction before it can go into effect.

Alternatively, if a large financial institution seeks a control relationship, it can structure its investment as a majority investment or an M&A transaction. Some advantages of a majority investment include: (a) providing more investor control over the AI fintech company than in a minority investment; (b) allowing the investor the opportunity to enhance the operational efficiency of the AI fintech company and address any existing risks (e.g., amend existing material agreements to address deficiencies); and (c) providing the AI fintech company with a greater opportunity to leverage the infrastructure and expertise of the investor. Disadvantages of a majority investment include (w) subjecting the AI fintech company to regulatory oversight; (x) requiring a much larger resource commitment from the investor, which entails a higher level of risk, necessitating a much higher level of due diligence (query whether it may be more advantageous to acquire the entire AI fintech company); (y) requiring a higher level of investor responsibility and oversight with respect to the operations of the AI fintech company, including regulatory compliance; and (z) integration issues with respect to the cultures of the investor and AI fintech company. The effect of the investor obtaining control of the AI fintech company is one of the most important factors for the investor's consideration. Generally, majority investments require a much more thorough due diligence investigation of the company than minority investments. The investor will need to assess the AI fintech company's current operations and marketing strategies (including the AI fintech company's website) and review its contracts, in each case with a particular focus on data security and regulatory compliance, as discussed more fully below.

41 Note that a potential alternative path for a bank holding company that has elected "financial holding company" status to invest in AI fintech companies is under the merchant banking authority in section 4(k)(4)(H) of the Bank Holding Company Act. This article will not attempt to address merchant banking authority, in part because its requirements (including with respect to the "routine management or operation" of a merchant banking portfolio company) are relatively restrictive.

In extreme cases, it may be necessary to shut the AI fintech company down for a period of time to resolve any major issues identified in due diligence.

Lastly, a large financial institution may wish to acquire full ownership of an AI fintech company in an M&A transaction. Each of the advantages and disadvantages of a majority acquisition apply to an M&A transaction, often to a greater extent. A key additional advantage of an M&A transaction is the flexibility provided, more specifically the opportunity to utilize a number of different structures to address specific risks (e.g., the use of an asset sale to protect against pre-closing liabilities). Some key disadvantages of M&A transactions include (a) requiring the highest level of due diligence and (b) concerns related to retention of key employees are at their peak.

The buyer's due diligence of an AI fintech company in an M&A transaction should include a confirmation of ownership of intellectual property and software, a personnel assessment and an evaluation of regulatory and data privacy risks. Analyzing the source code underlying the IP is critical. Open source code licenses may require disclosure to the public domain of all or a portion of the source code into which the open source code subject to any such license was incorporated. To reduce its risk, the M&A buyer should also seek to negotiate strong seller representations in the transaction documents with respect to matters such as ownership of IP, outbound licenses of the IP, use of open source code, the formatting of the source code (i.e., that it has been documented in a manner that enables a programmer of reasonable competence to understand it, manipulate it, etc.), compliance with data protection laws and best practices, and other similar matters.

The buyer of an AI fintech company should also seek to address due diligence issues and risks that are particular to AI providers. For example, the buyer should include compliance with law representations and covenants that allocate strict liability to the seller for machine learning output regardless of whether any breach is "intentional" or "negligent" or is known by the seller. Particularly where the AI fintech company engages in lending or making underwriting decisions, the buyer should address liability for discrimination and fair lending compliance, including for any disparate impact. The buyer may also seek a representation that decisioning criteria are "explainable" or at least diligence the design criteria of the AI fintech company for explainability. Cybersecurity and data privacy representations and covenants may also need to be augmented in light of data-intensive AI systems.

The buyer may seek to impose covenants in an M&A transaction that obligate the AI fintech company to address certain issues prior to closing, such as requiring the AI fintech company to bring its operations into compliance with data protection laws (including implementing any necessary changes to its IT systems), engaging a consultant to undertake a review of open source code, making changes to its marketing materials, obtaining any additional state or third-party licenses to operate the business, or renegotiating or terminating certain problematic contracts. Depending on the M&A buyer's leverage, it should also consider including closing conditions related to these matters to avoid being forced to close the acquisition and make these changes itself post-closing, which shifts the risks associated with any necessary shutdown to the buyer.

Lastly, as part of its due diligence process, the M&A buyer should identify key employees to retain following the closing. As mentioned above, there may be substantial differences between the cultures of the financial institutions buyer and the AI fintech company. Employees will often be moving from a relatively autonomous position with modernized infrastructure at the AI fintech company to a much more structured environment, often with restrictive

and outdated legacy infrastructure, at the buyer. Considering the importance of key employees, such as lead software engineers, to the AI fintech company, the buyer should ensure it is offering attractive compensation packages to encourage these employees to remain following the closing.

Conclusion

As shown in our discussion above, transactions involving investments in AI include a wide spectrum of possible structures, with legal and business issues that vary based on the transaction type. Financial institution investors should first define their AI goals and strategy, and then attempt to align their investment tactics with their AI strategy. As these AI strategies evolve, so will the transactions for investing in AI.

Smart Board Leval Questions to Ask About AI

This article first appeared on Directors & Boards website.

Rebecca S. Eisner Brad L. Peterson

Artificial intelligence, or "AI," raises legal and ethical issues beyond those generally found in investments in technology. Due to the rapid growth in this area, the lack of standards for evaluation and oversight and the risks associated with AI use, AI projects would particularly benefit from board inquiry and oversight.

Board members should ask the following questions as their company evaluates its use of AI.

Will AI Be Replacing Human Judgment?

As board members well know, our legal system relies fundamentally on human judgment in the areas of greatest importance. No board would simply turn over the question of whether a buyout offer is in the best interests of shareholders to an AI system, for example. Each board needs to inquire about whether sufficient consideration has been given to the potential uses of AI, particularly for businesses where legal compliance, fairness and adapting to new situations are important.

AI-based decisions must satisfy the laws and regulations that apply to your business. Of particular concern that AI-based decisions may discriminate because they rely on data that reflects a discriminatory past or looks only at correlation instead of causal factors. Companies that use AI tools in hiring, for example, need to ensure that these tools do not discriminate against certain protected classes of applicants or employees. In regulated areas like insurance, AI tools used for underwriting decisions will have to follow recently-issued requirements from the New York Department of Financial Services on the use of "unconventional sources or types of external data" to address the risk of unlawful discrimination and a lack of data transparency.

Companies can mitigate these AI risks by utilizing oversight, risk management and controls to meet legal compliance and ethical objectives. Data scientists who understand the AI tools and the context of the data and who implement controls designed to eliminate bias, inaccuracies and coincidence can reduce the chance of these unintended consequences.

In addition, AI systems will need to produce output that is transparent, auditable and that can be explained — sometimes called "Explainable AI." For the AI hiring tool example above, a company will need to be able to demonstrate that favorable hiring qualification scores of applicants are based on legitimate criteria, and not, on machine-determined prohibited factors such as race or gender identification.

What Are the Concerns Around the Data Used In AI?

Data is the fuel for AI. AI systems rely on statistical analysis and deliver the best results with large volumes of accurate, well-coded data. Companies using "machine learning" systems need a "data supply chain" to deliver a continued flow of current, accurate data.

Data use must comply with the privacy, data security, export control and other laws that apply to the data. For example, Europe now has tough data protection laws that prohibit the use of individual data for

automated processing to evaluate any behavior, preferences or location absent the explicit consent of the individual, and yet, automated processing of individual data to determine preferences is the hallmark of many AI tools. In addition, the data use must comply with any contractual requirements to data suppliers. These are often not well understood. To guard against these data pitfalls, board members should inquire as to the level of legal and regulatory diligence that has been done on the uses of data to fuel AI systems.

How Will the Company Protect What It Builds?

Patent, copyright, trade secret and other intellectual property (IP) laws were written to protect human creativity. IP laws in the United States do not square nicely with AI. Not only may your company not own AI that you pay to create, there may be no way to fully protect it under our IP laws.

Contractual and trade-secret protections are key elements of capturing and preserving value in the an investment in AI. These protections, to be effective, must be implemented before the AI effort begins.

How Will AI Be Implemented From a Contractual, Marketing and Operational Perspective?

IP protection may not be the only area where AI changes your business model. There may be effects on (and objections from) contracting parties, customers and employees. Recognize that your internal and external stakeholders have great (and possibly unrealistic) hopes for the benefits and, perhaps, also have considerable fears.

AI should be a cross-functional effort, including review and oversight by people focused on risk and potential harm. As a board member, you should inquire about the types of controls that are in place to avoid damage to relationships, brand, employees and communities.

How Will Evolving Laws Affect the AI Initiative?

There is an evolving understanding of how legal concepts such as reasonable care and agency will be applied to traditionally human processes now implemented by AI. There are also new laws related to AI, including "automated profiling," some of which carry substantial potential penalties. This analysis requires sophistication both in computing technologies and in the applicable laws generally, and you should probe for whether this level of analysis has been done.

How Does This Fit With General Risk Management?

The AI risk management framework should fit into the company's broader risk management framework and include standards for building, using and validating that AI models do not contain the problems discussed above. Company policies should require that new uses of AI undergo risk management review, and ultimately board review where appropriate. While it is vital to involve technical and security functions, we recommend that the board actively oversee whether the level of risk is appropriate for the company and whether the interests of internal and external stakeholders have been properly considered.

President Trump Launches AI Strategy for Federal Government

Rajesh De
Brad L. Peterson
David L. Beam
Kendall C. Burman
Alex C. Lakatos
Howard W. Waltzman

On February 11, 2019, President Trump signed an "Executive Order on Maintaining American Leadership in Artificial Intelligence" (the "Order") and, in doing so, set out a high-level strategy to strengthen the leadership position that the United States has maintained in AI. Important for companies, the Order sets off a number of opportunities for the private sector to give comments back to the federal government on how it can make changes that strengthen private sector AI development.

In recent years, China has made efforts to outpace the United States in developing AI technology, and, while not mentioning China specifically, the Order implicitly acknowledges the increased competition that the United States has faced from China in this area. AI is critical to US economic and national security interests, and the Order hopes to increase AI development through such measures as prioritizing AI research and opening up federal data to non-federal researchers.

The actions required by the Order are aimed at federal agencies that conduct foundational AI R&D, develop and deploy applications of AI technologies, provide educational grants, and regulate and provide guidance for applications of AI technology and will be coordinated through the National Science and Technology Council. While the Order does not place any obligation on the private sector, a number of the federal government activities will have an impact on industry. These include:

- **Increasing Access to Data:** The Order instructs all agencies to enhance private sector access to federal data, as well as to improve its quality and usability, for the benefit of the research community while protecting safety, security, privacy and confidentiality. The Order kicks off a number of steps that the federal government must take in order to achieve this, including publishing a *Federal Register* notice by which the public will be asked to "identify additional requests for access or quality improvements for federal data and models that would improve AI R&D and testing"; investigating the barriers to access or quality limitations of federal data; and updating implementation guidance for Enterprise Data Inventories and Source Code Inventories. In taking these steps, certain agencies must "identify barriers to, or requirements associated with, increased access to and use of such data and models," which include, among other things, privacy and civil liberty protections and the need for interoperable and machine-readable data formats. Making federal data more available to the private sector may also have implications for consumer privacy and is meaningful in the context of current legislative debates over comprehensive consumer privacy legislation. Additionally, the Order requires that the General Services Administration and other select agencies report back to the president on how to better enable the use of cloud computing resources needed to build AI systems.
- **Regulatory Review and Standards Development:** Within six months, the Office of Management and Budget ("OMB"), along with the participation of other relevant agencies, must issue a memorandum that instructs agencies on the

"development of regulatory and non-regulatory approaches…regarding technologies and industrial sectors that are either empowered or enabled by AI," as well as "ways to reduce barriers to the use of AI technologies." The public will be given the opportunity to comment on this memo before it's finalized. After the issuance of the memo, the agencies will then have six months to review their authorities affected by the memo and submit a plan to OMB on how they plan to achieve consistency with the memorandum. Separately, the Order also requires the National Institute of Standards and Technology ("NIST") to issue a plan within six months on how it will develop "technical standards and related tools in support of reliable, robust, and trustworthy systems that use AI technologies."

- **Prioritizing R&D:** The Order instructs agencies that perform or fund AI R&D to prioritize investment in AI R&D, although the Order is not explicit on how (or how much) AI R&D should be a priority. And, while the Order requires these agencies to identify which programs are AI R&D priorities, the Order does not augment the budget for an agency's AI R&D. It does, however, specifically instruct agencies to explore opportunities to collaborate with industry and other non-federal entities.
- **Workforce Development:** The Order instructs agencies that provide educational grants to consider AI a priority area in certain federal fellowship and service programs, these include alternative education and training programs and those that fund early-career university faculty who conduct AI R&D. The Order also requires the development of recommendations on STEM education regarding AI-related considerations.
- **Protecting National Security Interests:** The Order instructs agencies to develop an action plan to "protect the advantage of the United States in AI and technology critical to United States economic and national security interests against strategic competitors and foreign adversaries."

This Order marks a development in the Trump administration's AI policy, and, while it tackles a number of meaningful issues on how federal agencies should be organized around and should prioritize AI, there is much that the Order does not do, including increasing any funding for AI R&D, reforming or changing federal procurement of AI, or addressing important ethical questions on how AI should be developed and used. On the areas it does address, the Order leaves much of the specifics for further development by agencies. But it also includes important opportunities for private sector input. Specifically, public responses will be solicited regarding access or quality improvements for federal data and models to improve AI R&D, a draft memorandum issued by OMB regarding regulatory and non-regulatory approaches to AI and technical standards for AI technologies. Companies should evaluate whether they have interests that are affected by these developments and be prepared to offer applicable comments.

Who Owns Model Risk in an AI World?

This article first appeared on ABA Banking Journal website.

Reginald R. Goeke

Complicated computerized models and quantitative analyses are a fundamental mainstay in the financial services industry, from quantitative investment asset managers who use models to manage investment portfolios, to banks who use models to underwrite loans or monitor for money laundering or other behavior. With the benefits of those models comes several forms of risk, generally lumped together as "model risk."

Model risk generally refers to the potential for adverse consequences resulting from actions taken or decisions made based on incorrect or misused models or model outputs, and it includes risks related to errors in the quantification, coding or calculation process, use of improper or inaccurate data or other inputs, incorrect or inaccurate model design, or misuse or misapplication of models or model outputs. (The definition of a model "error" of "defect" is itself a subject of substantial debate, and often depends on the purpose and context for using the model. As noted in the article, whether a design decision rises to the category of "defect" will likely depend on the context of the use of the model, the model limitations disclosed to users, and the language of any agreement between the parties.)

The risk of such model errors is not theoretical. Over the past several years model errors have led to Securities and Exchange Commission enforcement actions, litigation and adverse headlines. For example, the SEC disciplined a quantitative investment adviser where an error in the computer code of the quantitative investment model eliminated one of the risk controls in the model, and where that error was concealed from advisory clients.

Similarly, where a robo-adviser advertised that its algorithms would monitor for wash sales but failed to accurately do so in 31 percent of the accounts so enrolled, the SEC found that the adviser had made false statements to its clients. Mortgage lenders have been accused of incorrectly denying loan modifications due to computer errors, and banks have suffered anti-money laundering compliance failures due to coding errors. As banks, asset managers and other financial services firms begin to deploy artificial intelligence or machine learning—whether in credit risk scoring, fraud detection, robo-advisory services, algorithmic trading, insurance underwriting or other areas—the potential model risks and related consequences increase.

Based on guidance from the Federal Reserve, the FDIC and other regulators, financial service firms have generally developed tools to identify, measure and manage those model risks. But that guidance predates the AI renaissance, and with the advance of big data, artificial intelligence and machine learning, potential model risks increase, and the controls needed to manage those risks and comply with regulatory and contractual obligations deserve additional attention.

For example, pursuant to the Federal Reserve's Guidance for Model Risk Management, the guiding principle of model risk management is effective challenge to the model, which requires critical analysis by objective, and informed parties who can identify model limitations and implement appropriate changes. Such effective challenge would include (among many other items) testing the theory and logic underlying the model design, validating the

model as well as integrity of data it uses, testing the performance of the model over a range of inputs, and implementing a governance model that permits independent review and assessment.

But in an AI world, when models work by identifying patterns in large data sets and making decisions based on those patterns, replication of the model's output (let alone reviewing performance across a range of inputs) becomes far more difficult. Further, when AI models apply machine learning to very large data sets, often from multiple sources, validating the integrity of such data becomes exponentially more challenging. And where model output may be generated in a black box based on the application of artificial intelligence, the ability of independent reviewers to effectively challenge any output becomes substantially more limited.

From a risk management and liability perspective, the questions that financial services firms should consider include, among others: How will a court determine (1) whether there were any defects in the model design, input or output; (2) whether any defect caused any adverse decision; (3) which party—among the model developer (or licensor), model user (or licensee), or the financial institution's customer—assumed the risk of the error or defect; and (4) the amount of any damages? These are the questions that courts and participants in the financial services industry will face in the coming years.

Is There a Defect in the Model?

When a bank or asset manager uses AI or machine learning and an adverse result arises—such as the poor performance of a loan or investment portfolio—the first question is whether the model was flawed in the first instance. Like human decision-makers, model-driven decisions may out-perform or under-perform relative to a benchmark and yet still be operating exactly as intended. In some instances, model defects may be objectively verifiable—such as the reference to incorrect cells or output in excel files, use of incorrect variables or the mis-specification of units. In other instances, particularly in the context of AI models, defects may be caused by a misinterpretation of underlying data, or reliance on coincidental correlations without causal connection, which may be much more difficult to detect. In still other instances, a model developer may make certain simplifying assumptions (e.g., disregarding data in a population set identified with ages over 120) that may impact on the model's performance. Such simplifying assumptions are a core part of "modeling" reality, and whether such assumptions cross a line into a "defect" or "error" may depend significantly on the representations made about the model and the context in which the model is intended to be used.

Given the challenges of explaining why any AI-driven decision was made, liability may often turn on the applicable standard of care (e.g., strict liability, negligence, etc.), the regulatory obligations of the model user (licensee), the types of representations made about the model, the known or foreseeable contexts in which the model may be used, and who (as between the plaintiff and defendant) bears the burden of proof. For example, an entity that touts that its models will monitor for wash sales but fails to fulfill that promise, may incur liability for the model's failure regardless of the source of any model defect.

A murkier issue may arise where a model developer markets its model as being able to reduce credit-related losses from portfolios approved using the model—but does not disclose that the model was tested using only populations from a certain geography or age. In that instance, if a financial institution using the model suffers substantial losses due to underperformance of the model with respect to populations for which the model was not tested, there will likely be substantial dispute as to whether the failure to test those populations constituted an error.

Did the Defect Cause the Adverse Outcome?

Assuming that a defect or error in a model can be demonstrated, it may still be an open question whether the defect actually affected a model's output. Many models (whether AI or not) will rely on multiple factors and rule sets. Even if an error existed in one part of a model, other portions may have corrected for the error, or may have led to the same result regardless of the error. To test for this, it may be possible to re-run a corrected version of the model with the same inputs, and thereby determine whether the error impacted on the model's output. In the context of AI models, though, which may use machine learning to detect patterns in millions of data points (e.g., credit application data, or asset management decisions), simply re-running the model with the same inputs may result in different outputs based on different machine learnings.

Thus, it becomes much more difficult to demonstrate whether or how any error affected model output. Although proof of causation is typically a plaintiff's burden, once a defect is demonstrated, some courts may implicitly shift the burden to the defendant to demonstrate that the defect did not have an adverse impact. In that event, an inability to explain (and show documentation of) the methodology and maintenance of the model (e.g., intended use, assumptions, theories, validations and testing, controls, versions) may limit an effective defense.
Who bears the risk of any model defect?

Even if a defect in a model caused an adverse outcome, potential legal claims will turn on which party assumed the risk of the model defect. This may turn on various tort, contract and similar legal principals, and depend on the relationships between the model developer/licensor (e.g., the party that develops and builds the model), the model user/licensee (e.g., the party that uses the model to make lending, investment, or other decisions), any customer of the user/licensee (such as a loan applicant), and any advisory client that invests in portfolios created by or managed with AI-enabled investment models. For example, where a credit card company uses an AI tool to build a better portfolio of loans, if there is a defect in the model that results in rejection of borrower applications, or that results in a pool of loans that underperforms expectations, who amongst the various entities will bear the risk for those decisions?

- **Model Developer versus Model User.** The liability as between a model developer and a model user is typically governed by the terms of an agreement, including representations, warranties and indemnification provisions. Some such agreements may be "as is" agreements, where warranty or indemnification obligations are disclaimed by the developer. In other instances, the model user may negotiate that the developer retains liability for its negligence or gross negligence. In that case, indemnification/warranty claims may turn on whether the developer/licensor applied industry-standard model controls (such as those outlined in the Federal Reserve's SR 11-7 Guidance), and the developer will need to be able to document its adherence to those controls. Further, liability may turn on the extent to which the model developer could reasonably foresee that the model would be used with certain populations or to make certain decisions. In many cases, liability allocation is likely to be heavily negotiated, subject to specific limited representations about model performance, and potentially subject to user representations about the use, testing and maintenance of the model.
- **Model User versus Affected Applicant.** Where a third-party customer (e.g., potential borrower) is denied credit based on the results of a potentially errant AI model, liability of the model user will likely turn on the user's compliance with various lending statutes, including ECOA, the Fair Housing

Act, FCRA, TILA and applicable regulatory loan origination and review requirements. Those requirements are beyond the scope of this article, but model users should conduct sufficient due diligence and testing with respect to any AI tool to understand and minimize the potential risks associated with use of the model, and should ensure that the model developer remains available to explain the model's performance to applicable regulators.

- **Model User versus Advisory Client.** In connection with investment portfolios constructed using an AI model, the contractual liability of the model user may turn on the extent to which model risk was disclosed to advisory clients and the extent to which the model user implemented model risk controls consistent with industry standards. As noted above, however, to the extent that AI models limit the effectiveness of traditional control processes (such as the ability to verify data quality, test model accuracy or challenge model output), model owners may be challenged to demonstrate compliance with standards that typically apply to model risk governance.

How Can Damages from AI Model Defects Be Quantified?

Assuming liability can be established, quantification of any damages still remains a challenge because a court would have to determine how the model would have performed absent any error or defect. For example, if an AI model has allocated assets improperly or created a loan portfolio with too much risk (based on the stated, intended purpose and usage of the model during the development stage), courts must first identify a relevant benchmark to determine how a portfolio might have performed absent any model error or defect.

For some models, it may be possible to correct the algorithm or coding and re-construct the portfolio absent the error. But where AI models are used to construct portfolios, and investment decisions depend in part of the assets already held by the portfolio—such as robo-adviser platforms—the iterative nature of the AI decision-making may make it difficult or impossible to re-estimate outcomes that would have existed but for the error. In a litigation context, plaintiffs may be given great latitude to argue about what actions might have been made or what outcomes might have occurred but for the error, with plaintiffs invariably seeking to apply a damage calculation methodology that results in the greatest amount of damages.

Potential actions for model developers and users. Given the additional complexities that AI models introduce for model developers and model users—including the "explainability" issues associated with AI models and the magnitude of data evaluated—those entities should consider steps to mitigate the liability risks. A few points of guidance emerge.

Model Developers

- **Curate Your Data.** Model developers should employ appropriate data curation controls. The adage of "garbage-in, garbage-out" is particularly applicable where the operations within an AI black box are difficult to evaluate. Developing a deep understanding of the sources of the data, triangulating the data with other available sources, and evaluating the data for potential bias are critical steps for developers to both take—and to document. In conducting this step, it is important that developers coordinate with legal and compliance, who understand the risks to be addressed and can help ensure that solutions are in a format that will be helpful when litigation ensues.
- **Improve Visibility Into Model Design.** Companies developing AI models should work with their model programmers to enhance the ability of reviewers to test and validate models. This includes

additional documentation regarding: the learning methods programmed into a model; the use of intermediate outputs that may help identify the data sets and decisions principally driving model outputs; and improved documentation of the quality assurance steps taken during model development and thereafter. Again, input from legal and compliance can help ensure that documentation is at a level that will be helpful in any future disputes.

- **Improve Contracting Steps.** Model developers/licensors and their counsel should clearly define the allocation of risk. Where possible, model developers may specify that agreements with model users/licensees expressly provide for the model in "as is" condition, and disclaim any implied warranties or indemnifications. Model developers should also be clear with licensees about any known limitations in models or data sources used to train those models.

Model Users/Licensees

- **Implement Meaningful Quality Control Procedures.** Model users/licensees acquiring AI models from third parties should implement meaningful quality control and due diligence procedures in the acquisition process. This would include a review of the data sources and the testing procedures used by model developers. Such diligence should inform the user's adoption of limits on the use of the model (e.g., using the model only to make decisions for populations similar to those from which the model was developed and tested). Such diligence should be coordinated with compliance and legal functions and documented for use in any future disputes.
- **Develop and Employ Effective Model Governance Processes.** The model users/licensees should adopt model governance policies and procedures to monitor the use of the model, and periodically confirm that the model's uses are consistent with the model's capabilities. Such governance models should include input by both technical staff and customer facing staff familiar with the ways in which the tool is being deployed and marketed. It should also include documented change-control processes, to be approved by all relevant stakeholders. Legal and compliance should ensure that disclosures and marketing materials are consistent with the capabilities of the model.
- **Include Human Input If Feasible.** Model users/licensees, where possible, should consider using models more for assistive intelligence, rather than as a pure decision-making tool. This would require employing personnel who can interpret the model outputs and, as necessary, apply their own judgment in making final decisions. Doing so can help ensure that questionable model decisions are identified earlier in the process and can provide an additional check to model decisions. Depending on the user's business model, human involvement in each model decision may not be realistic; but even in those cases periodic audits of model decisions can provide additional controls to the process.
- **Ensure Accurate Disclosures.** Model users/licensees should consider appropriate disclosures to customers, investors, and clients (including any individuals voluntarily using the AI-driven process) regarding the model's risks and limitations. Those disclosures should be reviewed both by compliance and legal functions, and also by the IT users of the model who are most familiar with the model's capabilities. Such disclosure may not eliminate liability, but where investors have the opportunity to make informed decisions after disclosure of the risks, the model user can more readily demonstrate that the investor assumed the risk of any error or defect in the model.

AI Legal Developments Related to Cybersecurity and Privacy

Kendall C. Burman Lisa Zivkovic
David A. Simon

Artificial Intelligence ("AI")[42] and machine learning[43] have recently been heralded as a near-panacea to a variety of economic and social problems involving everything from financial fraud and diagnosing cancer to public safety and workplace productivity. Yet privacy concerns have arisen in regards to key aspects underpinning AI applications, the opacity of algorithmic decision-making and the demand for sensitive personal information. A growing body of legislative and policy initiatives on both sides of the Altantic aim to protect against AI's potential dangers to individual privacy and security. In the European Union ("EU"), two key developments relate to the enactment of the General Data Protection Regulation ("GDPR") and the European Commission's ("EC") recent release of the "Ethics

Guidelines for Trustworthy AI."[44] A more narrow effort to address the harms of AI has been made by legislators in Washington state, who have, for the first time in the United States, proposed legislation that would impose obligations on organizations who use a particular form of AI to reach significant decisions about data subjects.

The EC's High-Level Expert Group on Artificial Intelligence ("AI HLEG") released its "Ethics Guidelines for Trustworthy AI" (the "Guidelines") on April 8, 2019 to provide stakeholders non-binding guidance on the ethical implementation of "Trustworthy AI," which involves embedding privacy protections into the AI system.[45] In consultation with various governmental, industry, and civil society stakeholders,[46] AI HLEG drafted the Guidelines to: 1) emphasize the

42 For the purposes of this article, the term "Artificial Intelligence" refers to "the theory and development of computer systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision-making, and translation between language." Oxford English Dictionary, https://en.oxforddictionaries.com/definition/artificial_intelligence.
43 Generally, we use this term to refer to the "field of study that gives computers the ability to learn without being explicitly programmed." Arthur Samuel, "Some Studies in Machine Learning Using the Game of Checkers," *IBM Journal of Research and Development* (1959).
44 The White House Executive Order on Maintaining American Leadership in Intelligence and the Norwegian Data Protection Authority's Artificial intelligence and privacy report are other examples of governmental initiatives focusing on the ethical implementation of AI. *See* Executive Order on Maintaining American Leadership in Intelligence (Feb. 11, 2019), https://www.whitehouse.gov/presidential-actions/executive-order-maintaining-american-leadership-artificial-intelligence/; Datailsynet, Artificial intelligence and privacy report (Jan. 2018).
45 High-Level Expert Group on Artificial Intelligence set up by European Commission, Ethics Guidelines for Trustworthy AI (April 8, 2019), https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines#Top.
46 AI HLEG members consist of professors across all disciplines, members of civil society organizations, such as the French Digital Counsel, German Research Centre for Artificial Intelligence, Fraunhofer Institute for Industrial Engineering, and DIGITALEUROPE, and individuals from the private sector, including companies like Orange, Element AI, Nokia Bell Labs, IBM, Santander Group, European DIGITAL SME Alliance, Bayer, AXA, and Google.

importance of a lawful,[47] ethical,[48] and a secure AI system;[49] 2) identify seven abstract requirements to realizing Trustworthy AI;[50] and 3) provide a "concrete and non-exhaustive" assessment list to operationalize the key requirements.[51] The AI HLEG recognizes that AI system implementation necessarily involves the collection of large data sets that likely contain personal and sensitive data, as well as obscure uses of such data.

The AI HLEG notes that privacy is a "fundamental right particularly affected by AI systems,"[52] identifying "Privacy and Data Governance" as one of the seven key requirements to realizing trustworthy AI. The Guidelines state that AI systems must "guarantee privacy and data protection throughout a system's entire lifecycle," from the collection of personal data to the generation of data about individuals. Although AI inputs can consist of sensitive and personal data, such as individual preference, age, gender, sexual orientation, or religious or political views, AI outputs can also make inferences using other data to independently generate such sensitive data. AI systems thus must establish individuals' trust in the gathering and processing phases by implementing external processes that reassure individuals that the data gathered and produced about them will not be used to "unlawfully or unfairly" discriminate against them.[53] The quality of data, or data that does not contain "socially constructed biases, errors and mistakes," must thus be ensured prior to training or feeding an AI system by qualified and properly trained personnel. The Guidelines suggest, as part of the assessment list to operationalize "Privacy and Data Governance," that stakeholders implement the following: mechanisms to flag privacy issues in data collection and processing; measures to secure data, such as encryption; and protocols and processes to train AI systems with minimal use of personal or sensitive data.[54] Although the discussion of privacy risks posed by AI is peripheral to the Guidelines' greater consideration of AI's impact on human rights, the ethical principles of human autonomy, fairness, and transparency that form the basis of AI HLEG's guidance for Trustworthy AI are also growing concerns for policmakers addressing consumer privacy regulations more generally.

Article 22 of the GDPR restricts automated decision-making and profiling, based on the automated processing of personal data, which produces legal effects for the data subject,55 unless the processing is: (1) necessary to enter into, or to perform, a contract; (2) based on explicit consent; or (3) authorized by national law.56 Automated decision-making based on the processing of personal data is of course at the heart of many AI applications. For these applications, the GDPR obligates covered organizations to: (1) provide the data subject "meaningful information about the logic involved"; (2) explain the signficance and the envisaged consequences of such processing for the data subject"[57]; and (3) provide the data subject with the right to obtain human intervention and context the

47 AI HLEG states that AI must adhere to the various "legally binding rules at European, national and international level [that] already apply or are relevant to the development, deployment of AI systems today,…[which include] EU primary law (the Treaties of the European Union and its Charter of Fundamental Rights), EU secondary law (such as the General Data Protection Regulation, the Product Liability Directive, the Regulation on the Free Flow of Non-Personal Data, anti-discrimination Directives, consumer law and Safety and Health at Work Directives), the UN Human Rights treaties and the Council of Europe conventions (such as the European Convention on Human Rights), and numerous EU Member State laws." *Id.*, at *6.

48 Trustworthy AI should adhere to four ethical principles: 1) respect for human autonomy; 2) prevention of harm; 3) fairness; and 4) explicability. *See id.*, at *12.

49 AI HLEG refers to security in the AI context as "technical robustness," which it defines as "including resilience to attack and security, fall back plan and general safety, accuracy, reliability, and reproducibility." *Id.*, at *14.

50 These seven requirements include: 1) human agency and oversight; 2) technical robustness and safety; 3) privacy and data governance; 4) transparency; 5) diversity, non-discrimination and fairness; 6) environmental and societal well-being; and 7) accountability.

51 *See id.*, at *24-35.

52 *Id.*, at *17.

53 *Id.*

54 *See id.*, at *28.

55 *See* Art. 22(1) of Regulation (EU) 2016/679.

56 *See* Art. 22(2) of Regulation (EU) 2016/679.

57 Art. 15(1)(h) of Regulation (EU) 2016/679.

decision.[58] Data processors and controllers under the GDPR thus cannot subject individuals to automated decision-making without explaining to the individuals the general processes that led to that decision and providing the option of human oversight. Article 5(1)(c) of the GDPR also imposes data minimization and retention limitation requirements. Personal data should be "adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed"[59] and "kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data will be processed."[60] In other words, data processors and controllers must restrict use of personal data to the amount and time necessary for a specific purpose. This provision of data minization can pose challenges to the implementation and oversignt of AI systems. AI systems require large data sets to be trained and the initial purpose for processing data can change based on what the machine learns. Humans also need access to data after decisions have been rendered to effectively provide oversight.

The Washington State Legislature proposed in its consumer privacy Senate Bill 2SSB 5376 ("WA Senate Bill 5376") a number of restrictions on the use and provision of facial recognition technology. Among those restrictions, is one that would require controllers using facial recognition to implement "meaningful human review prior to making final decisions where such final decisions produce legal effects . . . or similarly significant effects concerning consumers."[61] Processors that provide facial recognition services, on the other hand, would be required to explain how the technology works in clear and understandable terms, and prevent the use of such technology to discriminate against individual consumers under federal or state law.[62] While the WA Senate Bill 5376 is being considered by the legislature and is not actual law, these proposed restrictions on facial recognition technology provide further evidence on how policymakers are focused on the transparency, human agency, and fairness concerns raised by AI. Indeed, they accord well with the approach described above of the Guidelines, which emphasizes that stakeholders should implement mechanisms that facilitate explanation of the technical processes of AI systems to consumers in clear terms and promote human intervention to enhance equal treatment of consumers.[63]

AI stakeholders face important decisions over how to stay on the right side of the line with regard to both what the law requires as well as a growing body of best practices and ethical principles that apply to AI. Driving these decisions are core privacy, security, and ethical concerns that may be addressed through the implementation of mechanisms that facilitate data process mapping, anonymization, verification of consent, data quality, and human intervention as well as promote the security of personal data. Stakeholders should consider the role of human intervention in the decision cycle of the AI system, from the design and monitoring of the system to the ability to change a decision *ex-post* and have in place procedures that allow data subjects to exercise their rights. Furthermore, where the processing of data is based on explicit consent, one of the available legal grounds under Article 22 of the GDPR,[64] AI stakeholders must ensure that such consent meets the GDPR test (i.e. that it is freely given, informed,

58 *See* Art. 22(3) of Regulation (EU) 2016/679
59 Art. 5(1)(c) of Regulation (EU) 2016/679.
60 Art. 5(1)(e) of Regulation (EU) 2016/679.
61 Sec. 14(1) of WA Senate Bill 5376.
62 Section 14(2)-(3) of WA Senate Bill 5376.
63 Ethics Guidelines for Trustworthy AI, at *16, 18. Although the HLEG, GDPR, and the Washington Privacy Act emphasize the importance of human intervention as a safeguard to protect against AI's potential harms, HLEG has been criticized for overlooking the harm that can be done by humans. In other words, critics have argued that the system as a whole, which includes human intervention, must be transparent and both the system and the intervening humans should be accountable to consumers and end users for unfair or unlawful treatment. *See* Comments of the Center for Democracy & Technology on European Commission's High Level Expert Group on Artificial Intelligence (AI HLEG)'s Draft Ethics for Trustworthy AI, https://cdt.org/files/2019/02/comment - EU-Commission-HLEG-AI-guidelines-1.pdf.

64 Art. 22(2)(C) allows for automated decision-making where the data subject has given explicit consent.

specific and unambiguous) and implement mechanisms that would allow consent to be given, verified, and easily withdrawn. Stakeholders should also implement internal processes to ensure that data that is inputted and generated is non-discriminatory as well as favors anonymized data, thereby minimizing the use of personal or sensitive data. Finally, stakeholders must implement a cybersecurity program that ensures that the data is secure and not vulnerable to attacks.

To be certain, the legislative and policy landscape for AI is developing, but the efforts of HLEG, GDPR, and the WA Senate Bill 5376 , show how policy makers are wrestling with these important issues, and how trustworthiness, accountability, and ethics are equally important in considering the social impact of AI on data privacy and security.

Intellectual Property Rights in AI Data

Richard M. Assmus Brad L. Peterson

In many areas of research and development, businesses justifiably expect to be able to protect their innovations through intellectual property. What about when those innovations are in data? Data takes on heightened importance in artificial intelligence (AI) applications, where both the data needed to effectively train AI systems and AI output data may be have tremendous value. Here, we explore the availability of copyright and trade secret protection for data compilations under US law.

Copyright

The US Copyright Act protects original expression, not the underlying ideas or facts embodied in that expression.[65] Still, the US Copyright Act recognizes rights in compilations, which are defined as "a work formed by the collection and assembling of preexisting materials or of data that are selected, coordinated, or arranged in such a way that the resulting work as a whole constitutes an original work of authorship."[66]

Courts have grappled with the level of selection, coordination and arrangement required before finding original expression and, accordingly, granting copyright protection. Importantly, the underlying facts themselves need not be protectable for the compilation as a whole to be accorded protection. The most cited case on this question, as it relates to databases, is *Feist Publications, Inc. v. Rural Telephone Service Company, Inc.*,[67] in which the US Supreme Court reversed a ruling in favor of a phone book company against a competitor that had copied most of an entire phone book. The Supreme Court held that, in spite of the effort ("sweat of the brow") required to compile a phone book, the standard alphabetical listing of basic phone directory information was not sufficiently original to merit copyright protection. The Supreme Court noted, however, that "the originality requirement is not particularly stringent" and that "[p]resumably, the vast majority of compilations will pass this test."

Although *Feist* is often presented as the death knell for copyright protections in databases under US law, some cases applying the originality requirement soon after *Feist* actually found that particular data compilations merit protection, albeit narrow (see *Key Publications, Inc. v. Chinatown Today Publishing Enterprises Inc.* regarding a yellow pages directory and *Kregos v. Associated Press* regarding a baseball pitching form).[68] We have not, however, seen a case testing this proposition for a modern database in which data scientists made specific decisions about the selection, coordination or arrangement of the database or the particular data to compile for use in analysis. Nor has any case considered the creativity that may be involved in selecting a training data set for an AI system. A company that makes numerous choices with respect to the data that it uses to train the AI system—for example, by deciding to collect specific data fields and modify the training data to correct possible errors—may argue that its database should enjoy copyright protection, at least against large-scale verbatim copying. Certainly such a company would also benefit by documenting its

65 17 USC § 102(b).
66 17 USC § 101 (emphasis added).
67 499 U.S. 340 (1991).
68 945 F.2d 509 (2d Cir. 1991) and 3 F.3d 656 (2d Cir. 1993), respectively.

innovation process throughout, including any creative decisions made by the company.

In addition the models and output from AI tools may be creative works. However, these works are unlikely to be eligible for copyright protection if seen as machine output, as current US copyright law requires "an original work of authorship."[69] Although the definition of "author" is not fixed by the US Copyright Act, courts have found a human authorship as a requirement for copyright protection. In *Naruto v. Slater*, for example, the court required that a "person" or "human being" is required for authorship under the Copyright Act.[70] To achieve copyright protection, the company should have humans in any creative process using AI and documenting the human contribution to the work. The AI system may then be argued to be a tool, albeit a powerful tool, for humans to express human creativity in copyrightable works.

Trade Secret

The US Defend Trade Secrets Act defines "trade secret" as:

> "… all forms and types of financial, business, scientific, technical, economic, or engineering information, including patterns, plans, compilations, …, whether tangible or intangible, and whether or how stored, compiled, or memorialized physically, electronically, graphically, photographically, or in writing if –
>
> (a) the owner thereof has taken reasonable measures to keep such information secret; and
>
> (b) the information derives independent economic value, actual or potential, from not being generally known to, and not being readily ascertainable through proper means by, another person who can obtain economic value from the disclosure or use of the information."[71]

Trade secret protection is, thus, potentially available to companies leveraging secret data from their operations. Unlike copyright, trade secrets do not require an original act of authorship. Instead, trade secret protection requires the owner to take "reasonable measures to keep such information secret."[72] Like copyright, an intentional program of documented efforts to meet the legal standard will help to ensure protection.

Courts look to affirmative acts of the company claiming trade secret protection and may consider whether the company: (i) tracked database access and listed all employees, contractors, licensors, business partners or other third parties who could have misappropriated the company's compiled information; (ii) reviewed any agreements with employees, vendors, subcontractors and other service providers for confidentiality clauses, data security provisions, and restrictions on use of the data; and (iii) evaluated its database security measures and other internal confidentiality precautions at the start of the project and periodically thereafter.

To preserve trade secrets, companies should limit exposure to trade secrets (both within the company and outside of the organization) to those who need access. Companies can reduce the risk in any necessary access through confidentiality and other language in employee and independent contractor agreements. Courts often look for contract language

69 17 USC § 102(a).
70 *Naruto v. Slater*, 2016 WL 362231, at *1 (N.D. Cal. Jan. 28, 2016).
71 18 USCS § 1839(3) (emphasis added). In meaning, this definition is very similar to the Uniform Trade Secrets Act adopted in almost all US states.
72 18 USCS § 1839(3)(B)

between the parties to indicate that the disclosing company clearly communicated disclosure restrictions to the people receiving the trade secret and may consider non-disclosure agreements or non-compete language in services contracts as evidence in support of protecting a trade secret.[73]

Companies should perform a similar analysis with respect to vendors, subcontractors, data licensees and other agreements under which they allow another company to access data. Each third-party agreement that exposes a company's prospective trade secrets could threaten the company's claim for trade secret protection. In order to show that the company took "reasonable measures" to ensure secrecy, any such agreement should include clauses that require such reasonable measures from the licensee (such as a confidentiality clause) and specifically list the information to be protected (see *Events Media Network, Inc. v. Weather Channel Interactive, Inc.*, where a general restriction on disclosure of "Confidential Information" was not sufficient to show that the plaintiff intended licensed information to remain confidential).[74]

Finally, a company may be required to prove the reasonability of its security measures in order to make a successful claim for trade secret protection. Though this requirement is open to interpretation, courts agree with respect to a few best practices, including implementing password protections and restricting access to sensitive areas of facilities.[75] Of course, what is reasonable depends on the facts (the requirement is often stated as "reasonable under the circumstances"[76]). For example, companies that run AI systems may be required to prove spending on systems that comply with industry standards such as ISO 270XX and PCI DSS, implementing security protocols such as multi-factor authentication, and maintaining secure work environments for data scientists involved in AI projects.

Conclusion

AI is an emerging area where human creativity and control over secret data is a source of competitive advantage. However, copyright and trade secret laws were enacted before data became a substantial area of investment. To maximize the opportunities for legal protection, investors in data innovation should structure their projects with an eye to putting in place and continuing to maintain the best possible case for copyright and trade secret protections under the unique circumstances of the projects.

73 *Duggan v. Am. Family Mut. Ins. Co.*, 2010 WL 1268175 (E.D. Wis. Mar. 30, 2010).
74 *Events Media Network, Inc. v. Weather Channel Interactive, Inc.*, 2015 WL 457047, at *7 (D.N.J. Feb. 3, 2015).
75 Deepa Varadarajan, *The Trade Secret-Contract Interface*, 103 Iowa L. Rev. 1543, 1557 (2018).
76 *GSI Tech., Inc. v. United Memories, Inc.*, 2015 WL 1802616, at *4 (N.D. Cal. Apr. 20, 2015).

Getting AI Tools Litigation-Ready Is Crucial For Finance Cos.

Eric Evans
Alex Lakatos
Reginald Goeke

If your artificial intelligence tools are not litigation-ready, then discovery in a lawsuit contesting decisions those tools have made could quickly become a nightmare: Your company may suffer enormous distractions and decreased productivity as it struggles to address litigation requirements that are inconsistent with its AI systems, data and culture; may be subjected to onerous court orders that interfere with its ability to conduct its core businesses; may even suffer adverse judgments on claims that lack merit.

Preparing for disputes is a crucial consideration for any financial services company using AI and big data to make important decisions, such as whether to extend credit to potential borrowers or whether to flag a transaction as posing an anti-money laundering or fraud risk. As every decision-maker knows, once you start making decisions, you cannot please all the people all of time. Thus, litigation over AI decision-making is not a question of if, but when.

This article provides a road map for addressing discovery challenges intrinsic to AI, long before any lawsuits are filed, early enough that a thoughtful strategy and modest investment of resources can have a butterfly effect, multiplying to enormous value when disputes later arise.

Below, we first discuss why AI poses unique discovery challenges, different in quantity and quality for those arising from prior disputes over computerized models and decisions. Second, we discuss how savvy plaintiffs lawyers will seek to exploit those challenges to obtain strategic advantages in litigation — particularly in today's world, where some rules governing discovery and evidence still lag behind technological realities. Third, we set forth practical, actionable steps that financial services companies deploying AI can implement now, to help mitigate serious problems down the road.

We note at the outset that AI is a rapidly developing field and that most litigation over AI has yet to occur. Moreover, judicial efforts to grapple with the unique challenges AI poses are nascent, or even nonexistent. No doubt the future holds surprises. Our experience in other, related litigation contexts informs the article throughout; however, financial services companies will be best served by taking a flexible, nimble approach toward applying the recommendations below.

Why AI Poses Unique Discovery Challenges

Machine learning systems pose potential discovery difficulties beyond those typical for conventional algorithms or computer programs because AI systems are different in three critical respects: (1) inputs, (2) processing and (3) outputs.

Inputs

Andrew Ng, former chief data scientist at Baidu and an often cited AI expert, analogizes deep learning models to rocket engines that requires loads of fuel that is data. Machine learning studies and learns from data: It is "trained" on data. That thirst for data leads to several discovery challenges.

Data Volume

"Most applications of artificial intelligence require huge volumes of data in order to learn and make intelligent decisions." [77] Moreover, as algorithms become more sophisticated, they require even greater amounts of data. If a linear algorithm — a comparatively simple approach to machine learning — "achieves good performance with hundreds of examples per class, a nonlinear algorithm may need thousands of examples per class."[78]

Indeed, AI often functions by analyzing all the data that is available, e.g., reviewing all transactions, customer data, behavioral data and the like to spot money laundering risks or to assess creditworthiness. Producing and reviewing this data, as litigation often requires, poses significant challenges.

Data Sensitivity

In many instances involving financial services companies, the data used to train the AI will be sensitive. The data may include personally identifiable information, such as social security numbers and date of birth. It may reveal an individual's financial health and personal spending habits. It may contain medical information, such as spending on health professionals. In some cases, the financial institutions may owe duties of confidentiality to their customers. In other cases, while no official obligation may exist (or while obligations may be subordinate to discovery production obligations), the financial institution may still wish to protect its customers' privacy, whether for reputational reasons or as a matter of its own corporate values. This too, creates a challenge, especially when, as discussed above, vast troves of data are at issue.

Data Evolution

Machine learning systems may be designed to learn iteratively, refining their decision- making every time they receive additional data. A machine learning system that recommended extending credit on day one might make a different recommendation on day two, based on the system having seen more data, and having learned more, and having refined its internal model, in the interim. This presents discovery challenges for data. For example, is it even possible to go back and identify the data that the machine learning system trained on at a particular moment?

Data Retention

Many AI systems overwrite training data to conserve storage and other resources. Given the vast volume of data, and the fact data often ages out of usefulness, it may be impractical to maintain the data that led to a decision. But litigation-related preservation obligations do not automatically take practicality into account. Determining which data may be overwritten, when and how it might be preserved, and ensuring that the space exists to preserve it, can pose a significant challenge.

Processing

The manner in which AI tools analyze data to reach decisions is, more than any other factor, what separates AI from prior, algorithmic decision making programs. Those differences, however, create a host of discovery challenges.

Alex C. Lakatos, Eric B. Evans and Reginald R. Goeke are partners at Mayer Brown LLP.

77 Artificial Intelligence and Privacy, Datatilsynet (Norwegian Data Protection Authority) at page 4 (January 2018), available at https://www.datatilsynet.no/globalassets/global/english/ai-and-privacy.pdf.

78 https://machinelearningmastery.com/much-training-data-required-machine-learning/. Non-linear algorithms like a random forest or an artificial neural network are more sophisticated approaches to machine learning.

The "Black Box"

Before modern machine learning, algorithms made decisions that were rules-based, and so could be understood by studying the rules (the computer code) behind them. For example, a rules-based program might have a rule that provides that if borrower has a certain debt- to-income ratio above a certain amount, then the lender will not extend any additional credit to that particular borrower. Machine learning lacks such deterministic rules. Machine learning is probabilistic and uses statistical models. Machine learning might approach that problem above by building a model to answer the question: how much does this potential borrower resemble those who have paid-as-agreed, versus those who have defaulted?

The more complex the machine learning algorithm is, the more opaque the model and the harder it is to know why it made the decisions it made — e.g., what factors it weighed, how much weight it gave those factors, and how those factors interrelated. For example, machine learning easily beats human grand master chess champions, making moves that would not occur to them, based on reasoning they cannot fathom. Machine learning tools don't "think" like people: They have an "often quirky imagination."[79]

Output from a machine learning system may offer only limited insight into what is happening inside the black box. Machine learning is only concerned with the specific outcomes that its engineers instruct it to care about. It therefore may take a path to get to an end point that humans would consider to be cheating, undesirable or otherwise inconsistent with their intentions.

Sometimes, the outputs may instantly reveal that the machine has taken an unanticipated path, such as the following real world examples: (1) a robotic arm trained to slide a block to a target position on a table instead achieved the goal by moving the table itself and (2) an artificial life simulation where survival required energy but giving birth had no energy cost, one species evolved a sedentary lifestyle that consisted mostly of mating to produce new children to devour.[80]

On the hand, sometimes it will be far less apparent from the output that something unanticipated is happening inside the black box, such as a case in which AI trained to classify skin lesions as potentially cancerous learned that lesions photographed next to a ruler were more likely to be malignant.[81]

Data Retention Within the Black Box

The probabilistic decision-making model that comprises an AI tool (the inner workings of the "black box") may change from time-to-time, or even iteratively, as the AI learns from new data and adjusts the model accordingly. The AI system, however, may not be configured to retain values that change or are overwritten as the AI learns. For example, for deep neural networks, the value of the weights in each node may not be something the system has any means to preserve as the system repeatedly readjusts and refines them.82

AI Development

AI systems that are of significance to an entire organization, if not mission critical, typically are customized by and for the organization, and may be fine-tuned to reflect specific data sources, customers, marketing strategies, products and the like. Systems

79 https://www.technologyreview.com/s/612898/ai-is-reinventing-the-way-we-invent/
80 https://boingboing.net/2018/11/12/local-optima-r-us.html
81 https://www.wired.com/story/when-bots-teach-themselves-to-cheat/

82 A node combines input from the data it receives with a set of weights, that either amplify or dampen that input, thereby assigning significance to inputs with regard to the task the algorithm is trying to learn; e.g. which input is most helpful is classifying data without error? https://skymind.ai/wiki/neural-network.

may become so complex that even their creators have trouble understanding them. Such systems may have rules of operations different from products offered on the mass market.

Moreover, such systems are often in a process of constant updating and revisions by software engineers and data scientists, who may be doing anything from experimenting with new techniques for analysis to tweaking the inputs or outputs. There may not be one static set of code to produce, but millions of lines, with hundreds or thousands of owners, in a constant state of flux.

In addition, older machine learning systems may be superseded by newer versions. Once a system is no longer in active use, it may be difficult to maintain it in a usable form. On the other side of the same coin, many new lines of business do not succeed. If anything, this may be more true when it comes to new AI ventures, as the technology is novel and complex, and the regulatory environment is unsure.

The challenges described above are exacerbated when an AI system is shelved, or heading toward the dustbin. At that point, the incentives to preserve the system in working order, and to maintain data and records about the system, are diminished. Yet just as incentives to maintain and preserve the system are waning, the risk of litigation about the system may be waxing.

The more heavily documented a complex system is (with changes, modifications and even basic functionality documented and explained in real time), the more feasible it will be at a later date to explain the system and have a production (e.g., of code and training data) that is meaningful. But computer scientists and software engineers, particular those in nimble fintech startups, may not have a strong culture of documentation, and may even consider documentation inconsistent with their "flash on insight" programming methodology.

And even if computer scientists and software engineers routinely document their AI systems, that documentation may be more geared toward the needs of other computer scientists, not of litigants who need to explain AI systems for litigation purposes. In a high paced development environment, and one where turnover of AI programmers (who are in high demand) is routine, it is not uncommon for system documentation to consistent of shorthand and over simplifications.[83] But in a discovery context, plaintiffs' counsel may take such shorthand or simplifications out of context, argue that programmers are trying hide their decisions, or worse.

The "Secret Sauce"

Institutions that utilize AI often consider the exact programming and training of the AI an important trade secret. Some organizations, especially vendors supplying machine learning systems to financial institutions, may even promote their AI as better than those of others for reasons they cannot share (their "secret sauce"). Balancing the desire to protect valuable intellectual property from plaintiffs lawyers' who are likely to demand maximum discovery — both to help prove their case and, sometimes to help coerce a settlement — is yet another discovery challenge.

83 Nobel Prize winning physicist Richard Feymann tells a story from his time working on the atomic bomb in which junior physicists (many later recognized as geniuses of our time) were working so feverishly on a computer problem — at the time, based on punch cards getting out of order—that they didn't have time to explain the situation to their supervisor, much less document. His boss turned around and walked out of the room rather than interfere with the problem solvers working on the ground. Over 50 years later, his insight into how front line programmers see the world still rings true.

Outputs

Size and Complexity

Some AI systems have outputs that are huge and complex, and that can take a long time to generate. An AI tool that looks for fraud, for example, may review and risk weight thousands of transactions per hour. Further, the output may not be user-friendly; it may require expertise in a particular system to understand the meaning of the AI output. And that meaning may change over time, as the metrics for scoring or the interface are amended to reflect ongoing developments.

The Panda Problem

Pandas are animal known for living comfortably in their native habitat, but doing poorly when transplanted to another environment. Similarly, AI outputs may be usable and comprehensible within the system surrounding the AI tool, but may be hard to export in a meaningful fashion outside that native software environment.

For example, the information that AI system outputs may be (1) stored in deep storage, so that it first must be moved to fast storage before it can be searched, collated and utilized, (2) stored in a proprietary format that is exotic, as opposed to commonly known formats, such as .xls or csv; (3) subject to search and review only using specialized tools that may only exist in-house, and that may be understood only by in-house engineers.

Which of the problems above presents the biggest risk for your AI system?

Depending on how you deploy your AI, some of the discovery risks above may present greater challenges than others.

Predictive Analytics Tools

In the case of predictive analytics tools, such as AI that performs credit scoring, dispute are likely to focus on how the tool was trained and how it made decisions — which implicates the "black box" issues discussed above. For example, how does the black box predict credit performance? To what extent does it rely upon impermissible information, such as factors that are closely correlated with race, gender or other protected characteristics? And to what extent did the training data include such proxy information? Another contentious area for this type of tool may be quality data, and the possibility that inaccurate data led to inaccurate scoring — which implicates the data input problems discussed above.

RegTech AI Tools

In the case of regulatory technology AI tools, such as those used to mine the company's own data to identify trends or regulatory violations (e.g., Bank Secrecy Act violations, patterns in customer complaints), some disputes may focus on what the financial services company knew, and when the financial services company knew it — which implicates the data output challenges described above.

Other disputes may involve plaintiffs attorneys who want to mine the data themselves, and look for new problems as fodder for new or expanded claims. To do that, they will want all of the data input and their own access to the tool— which implicates the data input and black box problems. The same is true of disputes over whether the financial institution could, and should, have spotted a problem that it failed to identify, such as a Ponzi scheme orchestrated by one of the financial institution's customers that injured civil plaintiffs suing the financial institution.

Portfolio Development AI Tools

In the case of portfolio development AI tools, designed to improve the performance of a pool of

assets (e.g., a loan portfolio, a hedge fund), disputes are likely to arise in cases of underperformance and focus on allegations that owner/developer of the AI misrepresented its features and capabilities. A party that wishes to demonstrate that the tool did not perform as advertised (and that fault lies with the tool developer) may seek to scrutinize the tools outputs, its training, development and inner workings — which implicates black box and data output issues.

An adverse party that wished to demonstrate that the tool worked well, and underperformance is due to misuse (and that fault lies with the tool operator) may seek the scrutinize the parameters that the tool user adjusted and the quality of the data that the tool user input — which implicates a different set of concerns more focused on inputs.

Marketing/Sales AI Tools

These tools, which may mine customer data to enhance customer service or to identify marketing and sales targets for particular product, or which may interact directly with customers (e.g., chatbots), may raise yet another set of challenges, e.g., data input related in the case of data mining.

How Plaintiffs Lawyers Will Seek to Exploit Discovery Challenges for Strategic Litigation Advantages

Below, we discuss several areas where the perfect storm of discovery jeopardy may arise from the intersection of (1) complex modern AI, (2) rules of civil procedure (and common law guidance) that in some respects may lag technological developments and (3) aggressive plaintiffs lawyers. In particular, we discuss issues relating to preservation, production, and proof.

Preservation of Documents

The Risk: Sanctions

"Spoliation is the destruction or significant alteration of evidence, or the failure to preserve property for another's use as evidence in pending or reasonably foreseeable litigation."[84] Generally, trial courts have broad discretion to impose sanctions for spoliation. Depending on the forum, sanctions may include, among other things, precluding a party from presenting evidence on topics addressed in evidence that was subject to spoliation, allowing evidentiary inferences that the missing evidence would have been adverse to the party that failed to preserve it, finding certain issues conclusively established against the party that failed to preserve evidence, and entering a default judgment against the party responsible the spoliation.[85] In determining what sanctions (if any) are appropriate, courts generally will consider whether the party's preservation efforts were reasonable and undertaken in good faith.

In 2015, in recognition of the complexities and difficulties of preserving electronically stored information and the overzealous application of sanctions in certain cases, the Federal Rules of Civil Procedure were updated to provide guardrails constraining the ability of federal courts to impose sanctions for spoliation of ESI.

Specifically, under Rule 37(e)(2), a court must conclude that a party's spoliation of ESI was intentional (and not merely negligent or grossly negligent) before imposing more serious discovery sanctions, such as an adverse inference or default judgment. This rule helps protect parties utilizing AI tools, although some courts have been willing to infer intent from the circumstances of the preservation failure itself,[86] and some courts have allowed the jury

84 *West v. Goodyear Tire & Rubber Co.*,167 F.3d 776, 779 (2d Cir. 1999).
85 See e.g., FRCP 37; NY CPLR § 3126.
86 See, e.g., *O'Berry v. Turner*, 2016 WL 1700403 (M.D. Ga. Apr. 27, 2016) (holding that failure to preserve ESI, reliance on a single hard

to decide in the first instance whether the failure to preserve ESI was intentional.[87] State courts also may allow discovery sanctions for the spoliation of ESI on a showing of negligence or recklessness.[88]

How Plaintiffs Attorneys Will Try to Exploit the Situation

As detailed above, preserving incoming data streams, data explaining the working of AI and data output, all pose challenges. Plaintiffs are likely to try to exploit these difficulties in various ways, including:

- Sending letters making unreasonable demands for document preservation at the outset of the litigation, and revisiting the issue during meet-and-confer discussions on electronic discovery;
- Seeking discovery of documents (such as manuals, code and code documentation) and taking depositions (e.g., of corporate representatives, of engineers) to test what preservation might be feasible;
- Employing "experts" who will take unreasonable and unrealistic positions about what preservation is possible;
- Exploiting a court's lack of understanding about AI, playing on the common misperception that preservation of computer data is largely a straightforward matter (e.g., precluding automatic overwriting of certain data), and by arguing that preservation is simple and inexpensive to achieve; and
- Asking the court to infer that any data not preserved was a product of bad faith and seeking discovery sanctions.

Production of Documents

The Risk: Sanctions

Generally, parties must produce relevant information as part of the discovery process, typically in response to discovery demands. The production obligation is not unlimited, however. In federal court, for example, documents requests must be proportional to the needs of the case.[89] Similarly, state courts generally permit objections to discovery requests that are unduly burdensome. In disputes over production of AI, however, there is little guidance over where the proportionality/unduly burdensome line should be drawn.

To the extent the parties cannot agree on what data about an AI system will be produced, the next step will be seek guidance from the court. Once the court has determined and entered an order governing the required scope of discovery, failure to comply may lead to discovery sanctions such as those described above.[90]

How Plaintiffs Attorneys Will Try to Exploit the Situation

As detailed above, producing training data, decision-making data, data describing the inner workings of an AI system, and AI system outputs all pose many challenges. Plaintiffs are likely to try to exploit these difficulties in various ways, including:

---

copy, and loss of that hard copysupported a finding of intent to deprive).

87 See, e.g., *Cahill v. Dart*, 2016 WL 7034139 (N.D. Ill. Dec. 2, 2016) (holding that the jury should make the decision whether prison officials had intentionally allowed a crucial party of a videotape segment to be overwritten).

88 See e.g., *Pegasus Aviation I, Inc. v. Varig Logistica S.A.*, 987 N.Y.S.2d 350 (1st Dep't 2014) (adverse inference instruction appropriate in cases of negligent spoliation); *Scott v. Garfield*, 912 N.E.2d 1000, 1007-09 (Mass. 2009) (same); *Roebuck & Co. v. Midcap*, 893 A.2d 542, 548-52 (Del. 2006) (adverse inference instruction appropriate in cases of reckless spoliation).

89 See FRCP 26(b)(1).

90 See, e.g., FRCP 37(b)(2).

- Making unduly broad discovery demands;
- Making discovery demands the seek sensitive customer information;
- Making discovery demands that seek trade secrets related to the functionality and operation of the financial services institution's AI tools;
- Arguing that the financial institution has possession, custody or control of information concerning AI tools that resides with the financial institution's third-party vendors, and that the vendors are reluctant or unwilling to provide to the financial institution;
- Arguing that any materials produced are inadequate;
- Requesting that their experts be afforded direct, onsite access to the financial service's AI systems;
- Seeking court orders requiring the financial institution to produce data that goes beyond what is practical, and perhaps even beyond what is possible; and
- Filing motions for sanctions.

Proof Based Upon AI-Generated Evidence

The Risk: Inability to Present Evidence to Support a Defense

Generally, to have an AI model, or output generated by an AI tool, admitted in evidence in a judicial proceeding, the party presenting the evidence must "authenticate" it; that is, the party must demonstrate the evidence is what it purports to be.[91] "Among the factors courts may apply in determining whether a proper foundation for admission of computer-generated evidence has been laid include whether the computer was standard and in good working order, whether the operators of the equipment were qualified, whether proper procedures were followed, whether reliable software was used, whether the program operated properly, and the exhibit derived from the computer."[92]

In his treatise on federal evidence, Judge Jack Weinstein explains the rigor required to authenticate computer-generated evidence will depend in several factors, including (1) the quality of the data input, (2) the complexity of the algorithm, (3) whether the problem is routine or novel, and (4) whether the output can be tested and verified.[93]

Recent amendments to the federal rules streamline the process for authenticating "a record generated by an electronic process or system that produces an accurate result." That rule is intended for routine computer-generated evidence, such as electronic phone log. By contrast, AI models with inputs, weights and outputs that are in flux, or that are novel and hard to comprehend, may encounter authentication challenges.

How Plaintiffs Attorneys Will Try to Exploit the Situation

As detailed above, authenticating AI models and outputs for admission into evidence may be challenging. Plaintiffs are likely to try to exploit these difficulties in various ways, including:

- Seeking discovery of all facts that may bear on authentication;

91 See Fed. R. Evid. 901(a) ("The requirement of of authentication or identification as a condition precedent to admissibility is satisfied by evidence sufficient to support a finding that the matter in question is what its proponent claims.")

92 https://www.americanbar.org/content/dam/aba/administrative/labor_law/meetings/20 11/ac2011/123.authcheckdam.pdf .

93 Jack B. Weinstein & Margaret A. Berger, Weinstein's Federal Evidence at § 900.06[3].

- Filing overreaching discovery requests on issues they maintain bear on authentication;
- Retaining experts to challenge the authenticity of evidence; and
- Filing motions in limine to exclude a financial institution's AI-related evidence.

Approaches to Mitigating Discovery and Evidentiary AI Risks

Explainability

The better you can explain your AI tools — what the tool considered and why, how the tool made its decision and why — the stronger your position will be in a litigation disputing its decision and the effect of its decision.

Explainability, in general terms, has three aspects:

- Transparency: easy identification of the important factors in the tool's operation;
- Interpretability: easy identification and explanation of how the tool weights those factors and derives them from its input data; and
- Provenance: easy identification of where input data originates and what the data contains.

We recommend a two part approach to defending the actions of AI tools. First, be prepared to frame the discussion in terms of decisions people made. In other words, AI users will want to prove that the business decisions are the policy choices made by company management and the choice of algorithms and parameters by data scientists and programmers based on those policy decisions. The AI is not a decision-maker but merely a mechanism for implementing those business decisions. To do this, ensure the following:

- The management team specifies to its data scientists, computer scientists, software engineers and technicians how the company wants to tool to work, recognizing that those specifications are in fact the business decisions;
- The company's e-discovery/information governance team specifies to data scientists, computer scientists, software engineers and technicians how the company wants to store and access input, outputs, change logs, models and the like;
- The company consults its legal and compliance teams on the points above;
- The company employs "AI sustainers" to continual test and modify the tool to keep it working as the management team and the e-discovery/information governance team intended; and
- The company employs "AI explainers" who know how explain the tool's results.

Second, based on the guidance described above, the company can include features in its AI tool that further support explainability, such as:

- Code that permits auditing and testing;
- Explainable AI; i.e. the cutting edge, and still nascent, techniques that are beginning to allow a window in the AI "black box";
- Extra documentation that explains how the AI works and what choices were made about its features and functionality, for the benefits of current in-house employees, later in-house employees and later retained experts; and
- Thoughtful decisions about what facts and data to preserve and which to overwrite.

Explainability will be invaluable when confronted with the problems of production, preservation and proof described above.

First, when it comes to preservation of documents and data, explainability affords the company several advantages: (1) the advance creation of relevant records, such as system documentation, (2) an understanding of which records are important, so that the AI tool and related policies and procedures can be designed to better preserve those records, and (3) positioning the company to defend its choices about what and how to fulfill its document preservation obligations.

Second, and relatedly, explainability aids with document production by helping to (1) ensure that key records are created and preserved, as described above, (2) prepare the company and that AI tool to export data in a comprehensible, portable format, (3) position the company to argue that its production is appropriate and to defend against overreaching or misguided demands for other information or access.

Third, explainability goes to the heart of authentication: the better the company can explain its tool, the better the company can demonstrate that the AI model and outputs are what they purport to be.

Looking beyond discovery, to the merits of the case — i.e., when the company must justify the decisions of its AI tool — explainability will once again inure to the company's benefit. AI systems that aren't designed for explainability are often difficult to defend. Plaintiffs will provide expert testimony stating that the AI should have resulted in a one set of decisions that, not surprisingly, establish that they were harmed. In response, defendants will proffer their own evidence, usually from an expert, that tries to show that the AI functioned as intended and plaintiffs were not harmed.

But this battle of experts is fundamentally biased toward plaintiffs. When their experts are not presented with an explainable AI, they can simply provide favorable assumptions to traverse any difficult-to-explain aspects of the AI. They can do this because, as a practical matter, once an AI becomes too complicated to explain elegantly, a finder of fact will default to the simpler and cleaner explanation of the messier and more complicated one. Further, teams that create AI tools that aren't optimized for explainability will often throw off statements about "fixing" a "broken" AI system that feed directly into plaintiffs' narratives and undercut defendants.

An AI system optimized for explainability, though, can become almost a witness in its own defense. Design documents written for an audience of regulators or end users will make defendants' points better than design documents written for doctors of computer science. Data retention decisions can provide the key data points required to demonstrate the operation of the system instead of leaving it a black box. Documented design meetings and reports from sustainers can give the defendant human-scale stories, in human language, describing the AI — as opposed to mountains of raw data and near-indecipherable source code. And that change can level the playing field for companies defending their business- critical AI systems.

Storage of AI Tools and Information

It is also important to consider, at the outset of an AI project, where the tool will reside. There are several advantages to on-site storage from a litigation-ready perspective. First, litigation may require decades-long retention of data stores, which can add up. Second, if you own a server, you can always turn it off and physically shelve it, if necessary to preserve a legacy system.

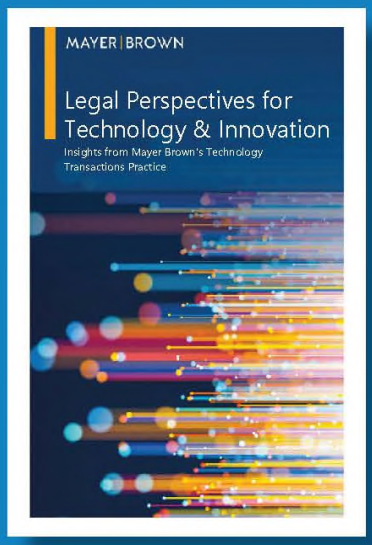
As for data, consider whether storage of higher risk data is necessary to your project. For example, if there are categories of personally identifiable information, health information or financial information that you do not need, consider whether you can avoid collecting and keeping that information. In the alternative, consider whether it is possible to anonymize information so that it's no longer personally identifiable.

Conclusion

Because AI tools are becoming ubiquitous in the financial services ecosystem, and because AI tools are more involved in decision-making than their predecessors, you should anticipate a flood of AI-related disputes. Companies that fail to prepare may find themselves drowning. Those that ensure that their AI is litigation-ready, by contrast, are well positioned to stay afloat. Now, when the levies have yet to break, is the time to act.

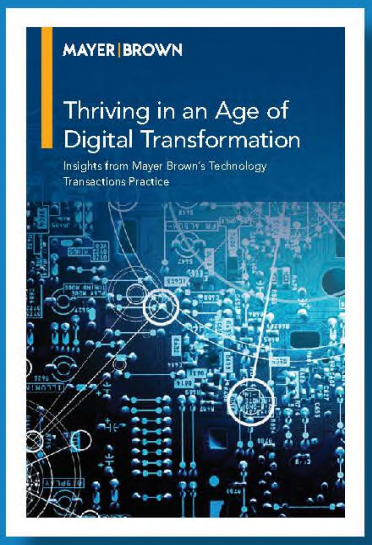
GET A COPY OF THE BOOK

Legal Perspectives for Technology & Innovation



GET A COPY OF THE BOOK

Thriving in an Age of Digital Transformation



VIEW TECH TALKS VIDEOS HERE

Mayer Brown's Tech Talks Series

VIEW THOUGHT LEADERSHIP FROM MAYER BROWN'S ARTIFICIAL INTELLIGENCE & FINANCIAL SERVICES SYMPOSIUM

Artificial Intelligence & Financial Services

VIEW REPORT/LEGAL UPDATE HERE

If Only: US Treasury Department Report Creates a Wish Tree of Financial Reform for Fintech

VIEW WEBINAR HERE

How Robotic Process Automation and Artificial Intelligence Are Changing Outsourcing in the Consumer Financial Services Industry

VIEW ADDITIONAL MAYER BROWN FINTECH PERSPECTIVES, LEGAL UPDATES AND THOUGHT LEADERSHIP

VIEW ADDITIONAL MAYER BROWN FINTECH EVENTS

LAW360

FIRM OF THE YEAR IN 2016, 2017 AND 2018

Practice Group of the Year:

APPELLATE: 2016 AND 2017

BANKING: 2016 AND 2017

CLASS ACTION: 2016 AND 2017

CONSUMER PROTECTION: 2017 AND 2018

INTELLECTUAL PROPERTY: 2017

TECHNOLOGY: 2016

CHAMBERS AND PARTNERS

GLOBAL

Technology Outsourcing: 2015-2018, Band 1

FinTech Legal, 2017

USA

Technology Outsourcing: 2004-2018, Band 1

Financial Services Regulation: Consumer Finance (Compliance & Litigation), 2018, Band 1

Capital Markets: Securitization, 2018, Band 1

The LEGAL 500

Fintech – Tier 2

Outsourcing – Tier 1

Structured Finance: Derivatives and Structured Products – Tier 1

Structured Finance: Securitization – Tier 1

IFLR

2017 *IFLR Americas*

Structured Finance and Securitization Team of the Year

2019 GlobalCapital Securitization Awards

2017 & 2019 *GlobalCapital Securitization Awards*

Best Overall Securitization Law Firm

Best Law Firm for ABS

ASIAN LEGAL BUSINESS

LAW FIRM OF THE YEAR 2018

Technology, Media and Telecommunications