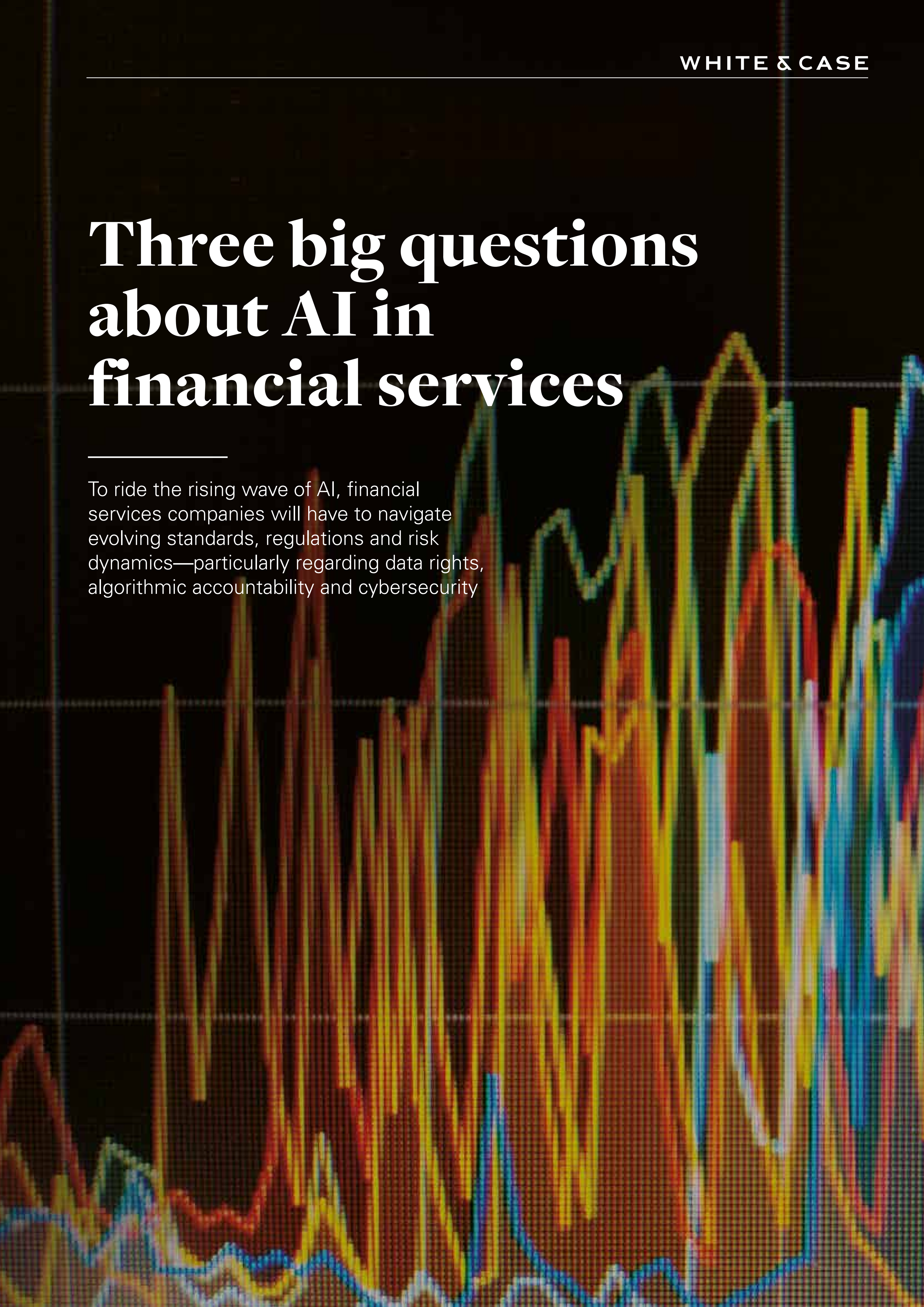
WHITE & CASE

Three big questions about AI in financial services

To ride the rising wave of AI, financial services companies will have to navigate evolving standards, regulations and risk dynamics—particularly regarding data rights, algorithmic accountability and cybersecurity



Three big questions about AI in financial services

To ride the rising wave of AI, financial services companies will have to navigate evolving standards, regulations and risk dynamics—particularly regarding data rights, algorithmic accountability and cybersecurity

By Kevin Petrasic, Ben Saul, George Paul and Steven Chabinsky

The success of artificial intelligence (AI) algorithms hinges on the ability to gain easy access to the right kind of data in sufficient volume. Put more simply, AI depends on good data. Even Google—which is famous for the pioneering work in AI that underpins its standard-setting search-based advertising business—makes no bones about the critical role of data in AI. Peter Norvig, Google's director of research, has said: "We don't have better algorithms, we just have more data."

Companies increasingly realize that data is critical to their success—and they are paying striking sums to acquire it. Microsoft's US$26 billion purchase of the enterprise social network LinkedIn is a prime example. But other technology companies are also seeking to acquire data-related assets, typically to acquire more than just identity-linked information from social media sources by focusing instead on vast troves of anonymized consumer data. Think, for example, of Oracle pursuing an M&A-led strategy for its Oracle Data Cloud data aggregation service, or IBM buying, within the past two years, both The Weather Company and Truven Health Analytics.

Early returns for companies making such investments are promising. Still, to unlock the full value of AI algorithms, companies must have access to large data sets, apply abundant data-processing power, and have the skills to interpret results strategically. Increasingly, those three elements are in the hands of the largest technology companies, fueling their market value. According to Kleiner Perkins Internet Trends Report 2017, seven of the 10 most valuable companies in the world are technology companies, compared with just three out of 10 in 2012.

Outside of tech, other industries are struggling to determine how AI can help power their future. But many financial services institutions start from a position of comparative advantage—they have large data sets and decades of experience using analytical tools, building models and employing large teams of software developers. More recently, they have also begun to incorporate data scientists into their ranks. Well positioned to leverage AI, financial services institutions have already begun to incorporate AI into parts of their business, such as algorithmic trading.

> “
> **To unlock the full value of AI algorithms, companies must have access to large data sets, apply abundant data-processing power, and have the skills to interpret results strategically.**

So how can financial services institutions best integrate AI into their operations—and, in turn, accelerate and improve their yield on AI investments? What we know is that, at the outset of their AI initiatives, companies will have to address three core questions: Who owns the data that is essential to AI? Who is responsible for AI decisions and actions? And what are AI's implications for cybersecurity? Answering these questions will often require grappling with complex issues that affect multiple stakeholders, sometimes with competing interests. And because the space is evolving rapidly, companies' answers are likely to evolve as well.

1 WHO OWNS THE DATA THAT IS ESSENTIAL TO AI?

Who owns this massive volume of data that companies are seeking to mine for insight? The answer to this question is an increasingly complex one, particularly given the explosion of unstructured data, the rise of partnerships that expand access to data, and the increasing ability to identify individuals when data sets are combined. Consumers, technology companies, third-party data providers, regulators and the panoply of institutions looking to realize AI's potential—all are stakeholders with complicated and competing interests in questions about who owns the data and for what purpose, what it's worth, and how it can rightfully be used.

Consumer privacy and consent

Consumers have a stake in which of their data is used, where and how it is used, and toward what ends it is used. But precisely what rights do they have? Legally, the answer may depend on where consumers live and what they agreed to when they gave access to their data. In the context of a global economy (and global data sets), organizations may often have to conduct complicated cross-border analyses and risk assessments, and resolve complex jurisdictional conflicts to determine the right way forward regarding consumer rights.

In response to this high level of complexity and competing demands, a variety of initiatives have been undertaken. The World Economic Forum, for example, launched a project, Rethinking Personal Data, that maps out what it describes as the "personal data ecosystem." The project calls for greater transparency into how data-driven polices are communicated, increased accountability for companies that harness data, and greater empowerment for individuals seeking to control how that personal data is used. But concrete results have yet to materialize.

Compounding the challenge, companies have begun to invest large sums to access consumer data, frequently through M&A. But, absent careful consideration, multiple deals, that may be layered on multiple and varied data-sharing arrangements, can sometimes muddy data provenance, increasing the risk that data will be duplicated and commingled. Separating commingled data can be extremely costly and difficult, if not impossible, because the data are often anonymized and aggregated into a pool.

This challenge may be even greater when it comes to unstructured data—such as text from emails, documents, social media postings, call-center transcripts and other sources. Unstructured data can be stored in relational databases alongside structured, transactional data, but extracting or associating it with structured data records can be difficult, not only as a matter of law, but also technically. AI can help with this, providing efficient ways to organize troves of unstructured data by adding and ordering metadata, essentially transforming them into structured formats.

> **Companies are working to ensure their data governance programs account for the explosion of unstructured data, which has become the feedstock for AI.**

AI and data governance

Financial services institutions have invested heavily over the years in data governance programs. They do so for a variety of reasons, including ensuring regulatory compliance; enabling data-driven decision-making; improving customer service; bolstering risk management; and addressing issues related to mergers, acquisitions and divestitures. But approaches to data governance can vary considerably from institution to institution. Some focus on data quality as part of data governance, whereas others view governance mainly as a set of management policies that are overseen by various councils. Still others consolidate everything, incorporating IT and related policies into one overarching data governance program.

But not all data in financial institutions is equal. Much of the data contains customers' personally identifiable information (PII), such as names, addresses, phone and social security numbers, information about their behavior regarding their personal finances, contractually protected data (e.g., information under a non-disclosure agreement), negotiated instruments and material nonpublic information.

Given these competing and complicated types and uses of data, financial services institutions are prioritizing the development of a cogent approach to data governance. Companies are working to ensure their data governance programs account for the explosion of unstructured data, which has become the feedstock for AI. This work includes identifying and mapping PII in structured and unstructured data, de-identifying and anonymizing customer information where necessary, and applying encryption and pseudonomization (a procedure by which the most identifying fields are replaced by one or more artificial identifiers) to comply with data privacy requirements.

Companies are reckoning with how AI affects their policies and processes. Banks, investment institutions and insurance companies are developing (or recognize the need to develop) policies that spell out AI-driven approaches to a variety of functions such as credit scoring and compliance, which increases processing speeds, minimizes labor and other resource costs, reduces human error and improves customer service.

Financial services institutions recognize that they can mine customer data to unearth opportunities for developing desirable, tailored products, such as providing customers with dashboards that aggregate and analyze data from disparate financial accounts and offering accounting services. Although banks enjoy a higher degree of customer trust than other industries, trust can erode if they do not protect customer data or provide third parties with unauthorized access to it.

Business partners

Partnering with other entities to collect, process and store customer data can further complicate questions of data ownership. Leasing data is also complicated. Data agreements are difficult to construct and interpret, and strict limitations on data usage can be difficult to enforce. Blockchain technologies may offer solutions by establishing data marketplaces that create auditable trails of provenance and ownership, logging appropriate use and attempted misuse.

The use of a third-party business partner does not relieve a financial institution of its obligations to comply with laws and regulations. A financial services institution retains responsibility for compliance, even when it outsources data processing and other functions.

Consider autonomous vehicles. Who owns the data about the car—the owner, the manufacturer or the software provider (if different from the manufacturer)? And if the same data is being used by multiple parties or algorithms simultaneously, how effectively can an individual ensure the data is used only for agreed-upon purposes?

Similar issues apply in the payments sphere. Multiple stakeholders—including consumers, merchants, payment providers and processors, and banks—may have interests in the data that is generated through payment transactions. When payments are made on mobile apps, the list of stakeholders might also include app developers, OS providers and wireless carriers. Sorting through who has which rights to what data—even when the customer has provided consent to its use—can be extremely difficult.

The answer hinges on data governance. Increasingly, financial services institutions must shift the emphasis from merely managing the life cycle of data to governing its use in context, including safeguarding data to ensure it is being used for its intended purpose—with the consent of the data subject and in compliance with global privacy laws, intellectual property laws, or other applicable laws, rules or regulations.

AI and Big Data have expanded the notion of business partners and data services. Data brokers have long traded personal and other data—such as age, gender and income data—in transactions largely comprising static lists and databases. Today any product vendor can potentially collect and provide access to data generated through use of its products.

The proliferation of the Internet of Things has only increased the number and type of products that generate data. Moreover, organizations increasingly use unstructured data that is available via real-time flows. It can be difficult to value data in real time, not to mention ensure that it is fit for its intended purpose. Real-time data use poses risks related to data provenance sovereignty as well as legal and regulatory compliance, whether an organization leases it or acquires it through M&A.

Valuing data in M&A

Mergers and acquisitions are commonplace in the financial services sector, and companies have become accustomed to the HR and IT issues involved. But one area that is far from agreed upon is how to value data in M&A deals—including the potential future value of data.

> “
> **Multiple stakeholders—including consumers, merchants, payment providers and processors, and banks—may have interests in the data that is generated through payment transactions.**

Consider a hypothetical situation in which Bank A buys Bank B, particularly because it wants to acquire B’s data. Bank A has its own trove of valuable data, as well as a deep understanding of AI and a staff that is able to combine the data sets to increase their aggregate value far beyond the sum of their parts. How can Bank B arrive at an appropriate valuation that ensures it gets a fair price for its data, given that the value of its data depends on factors that it may not understand or even be aware of? These and related data monetization, privacy and cybersecurity considerations are leading to an increase in the level of expertise, time and attention required during transactional due diligence to properly ascertain the value of (or liability embedded in) the data that is part of a transaction.

M&A deals also expose institutions to risks when acquirers do not fully comprehend what is in the data they are acquiring. During the 2007-2008 financial crisis, many data sets changed hands very quickly, and participants often did not have the time to sufficiently assess the risk the data may have presented. The situation is even more complex today given the emergence of powerful AI and deep learning algorithms that were not in use at the time of the financial crisis. In the future, such technologies will be even more common—and they will also be an important part of drawing up detailed risk profiles for data belonging to potential M&A targets.

2 WHO'S RESPONSIBLE FOR AI DECISIONS AND ACTIONS?

As AI becomes more sophisticated, companies will increasingly automate complex processes, including those that involve decision-making that requires judgment informed by experience. AI that incorporates machine learning and deep learning can leverage what it learns in order to write, in essence, new algorithms autonomously—and these new algorithms will affect the future decision-making and actions of this evolved AI.

Understanding what an AI program has learned and why it does what it does can be incredibly difficult. And because AI algorithms have an increasing ability to act independently, it may become difficult to assign responsibility to humans for decisions or actions AI takes. In some cases, existing laws and regulations clearly apply, but the grey areas are increasing, and anticipating the types of complications that will emerge as the scope and scale of AI-driven automation expands can be very difficult.

> “
>
> **Because AI algorithms have an increasing ability to act independently, it may become difficult to assign responsibility to humans for decisions or actions AI takes.**

Bias

In financial services—particularly in consumer finance—decisions about individuals' creditworthiness have traditionally been made using a transparent process with defined rules and relatively limited data sets. This transparency, however, may not always be achievable when AI drives big data. As AI is incorporated into financial operations, institutions run the risk that their algorithms may inadvertently make biased decisions or take actions that discriminate against protected classes of people—leaving financial institutions accountable, even if the alleged discrimination is unintentional.

There are three primary sources of bias in the AI process: data, training and programming. Bias can inhere in the data used to train machine-learning models—for example, in the assumptions used to create data sets for such models. It can arise when the size and scope of a data set used

Algorithms may inadvertently make biased decisions or take actions that discriminate against protected classes of people—leaving financial institutions accountable, even if the alleged discrimination is unintentional.

to train models is insufficiently large, and thus give rise to conclusions that are misleading or erroneous. Data sampling can compound this problem, as a given sample might not be representative of the data set as a whole. Bias outcomes can result from algorithms that input data reflecting bias and perpetuate those biases in the output.

The training of algorithms using what is known as "supervised machine learning" is also susceptible to bias. Even expert human trainers can potentially skew an algorithm by providing inputs and interpreting algorithmic results in ways that unintentionally reflect personal bias.

Most data scientists and developers go to great lengths to create objective algorithms. But even the most careful expert trainers are subject to contextual influences—cultural, educational, geographic—that can affect the assumptions that inform machine-learning code and skew algorithmic results. Beyond this, as algorithms learn on their own, they can draw facially neutral connections (i.e., connections that do not appear to be discriminatory on their face) that, nevertheless, adversely impact protected classes.

This kind of unconscious bias poses an important challenge, because bias does not need to be intentional to land financial institutions in hot water. Courts and regulators are expanding their use of the "disparate impact" theory of liability, which focuses more on discriminatory effects of credit decisions and policies than on a financial institution's rationale or motivation. Under disparate impact theory, if bias is seen to result in an alleged discriminatory act, there is no need to show that discrimination was intentional to establish liability—only that the discrimination occurred.

Recourse

Many hope that AI—particularly deep learning—will enable automatic and algorithmic trading to evolve beyond a focus on containing costs to driving profits by informing trading and financial planning. For example, AI can facilitate the use of big data to inform decisions on opening customer accounts and extending credit to customers. But it may be difficult for financial institutions to accept the risks of AI if algorithms are not accountable for results.

First, unaccountable algorithms can undermine US and global data privacy principles and requirements, leading to fines, penalties and other regulatory actions. The Federal Trade Commission (FTC) Fair Information Practice Principles (FIPPs) provide guidelines for entities that collect and use personal information. The FIPPs include safeguards to ensure that information processing in manual and automated systems is fair and provides adequate privacy protection.

In the European Union (EU), under the Charter of Fundamental Rights of the European Union and the Treaty on the Functioning of the European Union, persons have a fundamental right to protection regarding the processing of their personal data. Under the EU's General Data Protection Regulations (GDPR), personal data must be processed in a fair and transparent manner, often with the consent of the data subject. Among other rights, EU data subjects may obtain confirmation of whether a controller is processing their personal information, the recipients (or categories of recipient) with whom their data is being shared, and the existence and significance of any automated decision-making.

Second, it will be perilous for financial institutions to negatively affect consumer access to financial services via algorithms without justification or recourse to review a negative outcome. This will warrant the attention of the FTC, the European Commission and other consumer watchdogs that will investigate suspect algorithms, training sets and the underlying data.

Third, the costs of building and maintaining AI-powered applications should include mechanisms to safeguard customer and consumer information and provide a process for consumers to protest an algorithmic denial of financial services. If consumers have no recourse to a negative outcome of an AI algorithm, they may not give their consent to apply their personal and financial data to an algorithm. Financial institutions must ensure that all data used by algorithms to determine customer service is accurate and valid and that there are appropriate technical controls, including encryption or pseudonymization, to safeguard personal and private information

AI-powered applications should include mechanisms to safeguard customer and consumer information and provide a process for consumers to protest an algorithmic denial of financial services.

“

Competition law policy will dictate that regulators evolve their analysis of tacit collusion to accommodate non-human decision makers.

while it is being processed and stored.

Machine-learning technology is powered by complex mathematics. Deep-learning technology seeks to create overlapping networks of algorithms that are analogous to a brain's neural networks. Researchers are a long way from knowing which data is processed where in an AI neural network. That said, work is underway to test neural networks and determine how data flows through various nodes and layers, which will ultimately help clarify how decisions are made.

Antitrust, collusion

AI could also have important ramifications for competition law, as pricing mechanisms shift from people, whose actions are more readily covered by existing competition law, to computer algorithms. Competition law has historically focused on corporate actors who are seen to be complicit in limiting or distorting competition, in particular through collusion—agreement and intent to engage in efforts such as price-fixing in an oligopoly or across competing products.

In an AI-driven landscape, new questions arise, such as: What constitutes collusion in an algorithm-dominated environment? What are the boundaries of legality and collusion? What are the antitrust liabilities for developers and users of algorithms? What technology can monitor and constrain AI? If human intelligence can be defined and simulated by a machine, is it possible to create moral, ethical and law-abiding algorithms?

Even when computer algorithms are proxies for market players, overt collusion remains a violation of competition laws. An agreement to utilize algorithms to coordinate prices or reduce competition remains illegal. It is more difficult to determine legality when the collusion is not explicit but may be tacit—such as when prices are coordinated but there was no explicit agreement to coordinate them among participants.

Suppose similar computer algorithms operated by different companies to promote an optimal pricing strategy in which each predicts another's reaction to fluctuating prices. Suppose further that the algorithms determine that the best way to maximize profits for their individual corporate owners is to each set the same optimum price for a particular product. Would the algorithms' concerted action amount to tacit collusion to fix prices?

Will lawmakers allow similar algorithms to run in a competitive industry until their corporate owners learn of the concerted effort to set an optimal price among competing algorithms? Perhaps lawmakers will redefine "agreement" and "intent" for computer actors and introduce proscriptive elements for algorithms and limit the big data available for analysis. At minimum, competition law policy will dictate that regulators evolve their analysis of tacit collusion to accommodate non-human decision makers and consider demanding more transparency into the decisions made by algorithms to permit monitoring and alerting mechanisms for potentially anticompetitive behavior.

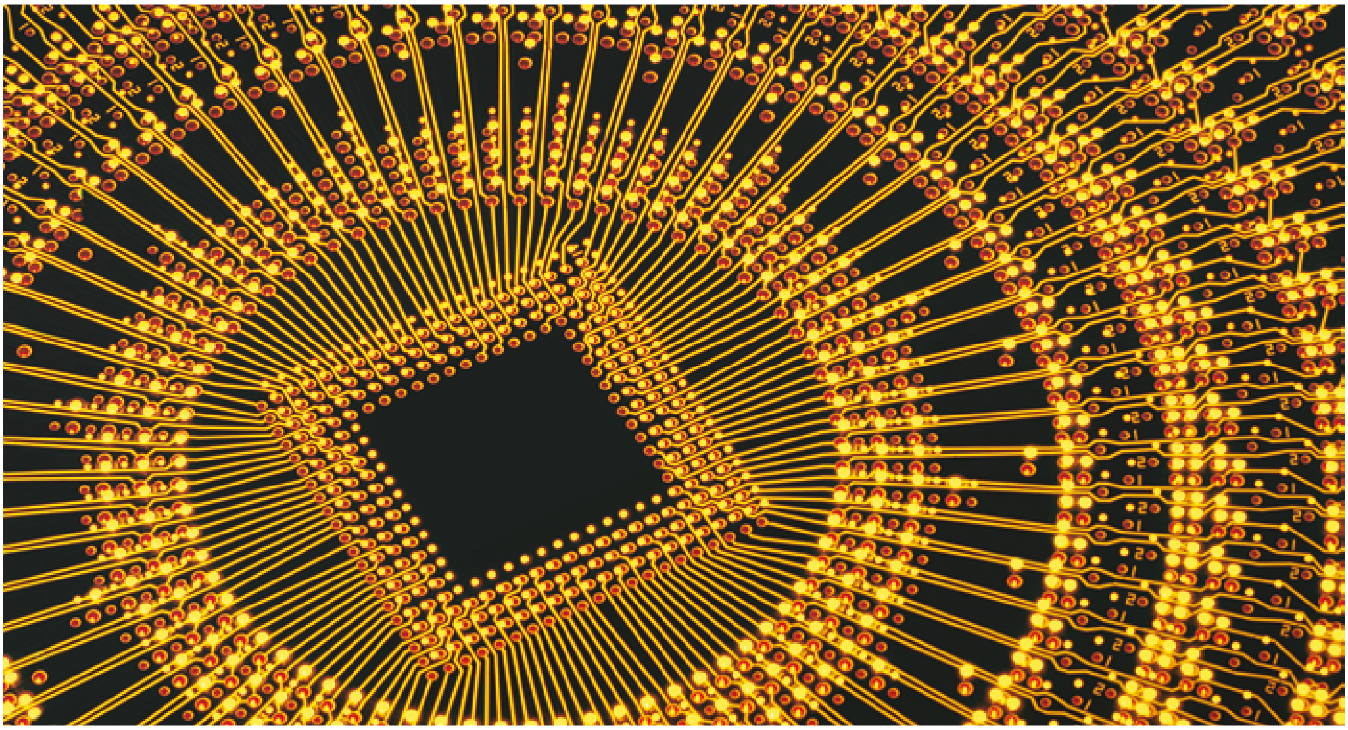
3 WHAT ARE AI'S IMPLICATIONS FOR CYBERSECURITY?

Organizations are already beginning to use AI to bolster cybersecurity by, for example, automating complex processes for detecting attacks and reacting to breaches. And these applications are becoming increasingly sophisticated as learning AI is deployed in the security context. But AI can open vulnerabilities as well, particularly when it depends on interfaces within and across organizations that inadvertently create opportunities for access by nefarious agents. And attackers are beginning to deploy AI, too, meaning they will increasingly develop automated hacks that are able to learn about the systems they target, and to identify vulnerabilities, on the fly.

White-hat defenses, partner vulnerabilities

Financial institutions engage in a broad range of activities that require them to collect, store and use sensitive information about customers, their finances and their behavior toward a variety of ends, from account creation to detecting criminal behavior in banking transactions. Sensitive data includes PII, while less sensitive but valuable information includes customer transaction data.

Besides information collected directly from customers, financial businesses also acquire data, including transaction data, from third parties and the internet, such as social network data, to inform AI-driven efforts to deepen their understanding of their customers. Each financial activity or line of business may use AI differently. Banks may use chatbots, for example, to improve customer experiences. Wealth planning and management services may use robo-advisors to add investment options to customer portfolios. Insurance companies may use AI in claims processing to improve workflows and identify fraud. Financial businesses might use AI to identify relationships that could give rise to



new services or identify fraud and money-laundering activities.

But commingling a variety of data in the service of AI can be fraught with risk. First, Big Data collection creates a growing attack surface that increases vulnerability to hackers, who look to breach security controls and access PII without authorization. Second, AI provides hackers with a tool to land and expand data breaches (although conversely, AI will also become essential in detecting and preventing breaches). Third, AI can help financial institutions identify new relationships, but in doing so it can sometimes recreate identities that have been masked to comply with privacy requirements. Fourth, commingled data makes it difficult to track whether a data set includes PII, who owns the data, and how the data can be used—in light of restrictions related to data subjects' consent and applicable laws and regulations.

Organizations that use commingled data are still required to know what data they have, where it is located and who owns it. Different laws and regulations apply to various data types across regions. For example, the GDPR will provide EU data subjects with rights regarding companies' processing and holding of their personal data regardless of the organization's location. Individuals in EU member countries can exercise these rights to identify the information collected about them and understand how it is used, as well as to access, review and demand the correction and deletion of that data under certain circumstances.

For compliance, financial institutions can track the provenance of data down to the record level and apply policy management and security controls, such as pseudonymizing and encrypting data. But the micromanagement of large data sets used in AI can be costly and require additional tools, such as applying metadata enhancement techniques to identify data under compliance requirements, tag it appropriately, and apply policy management to control data access and usage. Sometimes, businesses may destroy entire data sets rather than invest the time and resources to micromanage data provenance at the record level.

So-called "personal data vaults" have sprung up in recent years, enabling individuals to store personal data and manage its use, and even to grant anonymized access to it for a fee. Similar efforts have been undertaken with personal health data in recent years. So far, few if any of these efforts have proven particularly commercially successful, although health data is conceivably one area where personal data management could eventually proliferate.

> “
> **The large and increasing volume of data that financial institutions use for AI presents a large attack surface.**

Black-hat attacks

Much of the data used in the financial services sector is of great interest to hackers. The large and increasing volume of data that financial institutions use for AI presents a large attack surface. The security challenge may be compounded by partnerships with third-party data providers, particularly when the third parties provide real-time data via API and may not have the hardened, secure interfaces found in financial services.

Financial services institutions are generally required to conduct risk assessments of security measures and design an information security program to protect nonpublic personal data and sensitive business information. Financial institutions are also increasingly responsible for the security measures used by third-party partners and affiliates; from a regulatory and compliance standpoint, insured depository institutions are jointly accountable with third-party vendors.

The repercussions for things like data breaches and allowing unauthorized access to PII go beyond regulatory fines and penalties. Following a data breach, the erosion of customer trust and the tarnishing of a financial brand could shake—if not shatter—a company's ability to compete in the financial sector, where the secure handling of financial data is imperative. On the other hand, failing to take advantage of new technologies, including AI, is just as certain to leave a company trailing others currently poised to disrupt the financial sector.

The key is to move forward, but smartly. Fortunately, the cybersecurity industry itself is taking advantage of AI, leading to automated detection and response capabilities that were unheard of only a few years ago.

REALITY CHECKS

The challenges of implementing AI are real but the benefits are great—including increased speed and efficiency, reduced labor and resource costs, reduced human error, the ability to tailor products and services and otherwise improve the customer experience, and improved security. Financial institutions must proceed with care in the AI era, but few, if any, can afford to sit back or ignore AI.

To stay focused amidst the welter of activity in this space, companies should adopt three broad guidelines:

Set out clear principles and document strategies and processes

Any data set may contain the seeds of bias, and learning machines may take unwanted actions that no company can anticipate. Companies that articulate their objectives and show that they have a made concerted effort to comply with regulations and respect consumers can put themselves in a position of strength when challenges arise. This includes rigorously testing systems, and analyzing and documenting their performance. Those that do not do this will leave themselves vulnerable to reputational and legal complications, even when they have done their best to meet high standards.

Manage technology with technology

Machines will increasingly be used to monitor other machines, whether that involves regulatory technology (regtech) applications that enable companies to automate compliance or cybersecurity applications that enable companies to identify potential vulnerabilities or trigger responses to actual breaches. AI can enable these kinds of meta-technologies to learn on the fly so they can respond effectively even as the machines they are designed to oversee or defend against evolve in real time. Technology will become increasingly important as a means of managing technological complexity as data flows continue their exponential growth trajectories, and increasingly sophisticated AI goes mainstream.

Keep people front and center

AI programs are increasingly able to make sophisticated judgments and decisions and even write their algorithms based on insights they learned through experience to guide their future choices and actions. But most experts agree that we are a long way from a future in which computers can operate autonomously in contexts requiring sophisticated, dynamic use of judgment. It is always critical to remember that technology is a tool, and humans must oversee the machines they deploy to ensure choices made by algorithms make sense and align with social principles and regulatory rules. In the vast majority of cases, there is no substitute for human judgment. Companies that cede too much control to technology will increase their risk exposure significantly, if not catastrophically.

We are at the beginning of a journey that will take some time to unfold. The financial services sector will be one of the most important sectors forging new commercial applications for AI in the coming years, but it may also be among the most vulnerable to AI solutions that go sideways. Companies that balance the various interests of stakeholders will produce significant benefits for business and consumers.

> **“Technology is a tool, and humans must oversee the machines they deploy to ensure choices made by algorithms make sense and align with social principles and regulatory rules.**

WHITE & CASE

Kevin Petrasic
Partner, Washington, DC
T +1 202 626 3671
E kpetrasic@whitecase.com

Benjamin Saul
Partner, Washington, DC
T +1 202 626 3665
E bsaul@whitecase.com

George Paul
Partner, Washington, DC
T +1 202 626 3656
E gpaul@whitecase.com

Steven Chabinsky
Partner, New York
T +1 212 819 8718
E schabinsky@whitecase.com

whitecase.com

In this publication, White & Case means the international legal practice comprising White & Case LLP, a New York State registered limited liability partnership, White & Case LLP, a limited liability partnership incorporated under English law and all other affiliated partnerships, companies and entities.

This publication is prepared for the general information of our clients and other interested persons. It is not, and does not attempt to be, comprehensive in nature. Due to the general nature of its content, it should not be regarded as legal advice.