Intelligent Approaches to AI

The wide-reaching effects of artificial intelligence, and the potential unexpected consequences of its misuse, require a different set of questions for board oversight.

By Jesus B. Mantas

Despite the promise and increasing presence of artificial intelligence (AI) in everyday life—as well as the massive investment in it of more than $37 billion by companies this year, and the nearly $10 billion invested by venture capital firms in 2018—few boardrooms are prepared to oversee the risks and opportunities associated with AI.

Ensuring that companies make ethical decisions is clearly a boardroom matter. A new study by IBM's Institute for Business Value reported that more than half of the 1,250 executives surveyed believe AI actually can improve their companies' ethical decisions. Yet while 8 out of 10 directors believe the ethical questions raised by the deployment of AI are board-level issues, only 45 percent feel fully prepared to oversee them.

AI is different from other technology innovations, affecting how boards approach oversight. Boards are seeking concrete guidance on how to protect the interests of their companies and stakeholders impacted by AI. Specifically as it relates to the game-changing potential of AI, boards want clarity on what constitutes appropriate governance, how to anticipate and assess new types of risk, and how to effectively oversee management's definition of and adherence to AI ethical standards.

To keep pace with AI developments, directors should consider asking additional questions in the areas of business strategy, people management and culture, finance and controls, risk, and governance. Here's how.

AI's Unique Nature

AI's greatest strength also has the potential to introduce its greatest risk: AI's ability to learn. Unlike traditional software, where the machine follows a known set of instructions, an AI system is guided by algorithms that may continuously and autonomously adapt, refine, and alter its responses and decisions, so it may not always be apparent why an AI system reached a particular conclusion—and even if we agree with it, can we trust it?

An issue with an AI application that may seem innocuous in an everyday setting may have significant ramifications in a business context. For example, in a photo app, if the AI wrongly labels a face when grouping photos, it's a minor inconvenience. But if AI makes a biased judgment based on personal data and an insurance company then approves or rejects a claim or if a bank approves or rejects a credit application based on that judgment, the consequences can be significantly more severe.

AI performance is dependent on the same factors that affect how humans learn: what we are taught, how and by whom we are taught, and the logic we use to reach a conclusion. People educated in different ways with different sets of information may make drastically different decisions. The same can be said for the way AI is developed and trained, and how it applies the skill that it was created to perform.

To use AI in mission-critical applications, we need to be able to trust the decisions it makes. And in order to trust an AI system, there needs to be confidence in the data it is fed, the method used to train it, the rationale for how the system reaches its every conclusion, the ability to verify that those conclusions are aligned with the original intent of the algorithm, and that every step along the way is aligned with the company's purpose and ethical standards.

This means that when engaging an AI system to automate a task or to support decision making, it is essential to think through human involvement. Ensuring that there are humans looped into critical AI-supported tasks, and that these individuals represent and enforce the company's values and ethics, is similarly essential.

In addition, AI can muddy implications for who is at fault when something goes wrong. If traditional software goes askew because the programming has bugs, it's clear the software developer is responsible. When an AI application goes haywire, or when the outcomes and behaviors of the application are questionable, it may be difficult to determine where the fault lies within the algorithm, especially if multiple parties collaborated on the development and training of the application. Does fault lie with the data,

GETTY IMAGES

the process used to teach and evolve the system, or the people who trained it?

As in other fields, AI technology is developing faster than regulations. For boards, a question of paramount importance is, when does AI technology demand oversight to ensure that its strategic benefits and operational risks are clearly defined?

While many world and industry-specific bodies are rushing to provide guidelines, what's particularly challenging is the speed of AI's advancement and the scale of its adoption. Of specific interest to boards: the World Economic Forum (WEF) Global AI Council, organized by the Centre for the Fourth Industrial Revolution, is addressing the questions specific to AI and corporate governance. In early 2020, the WEF is expected to release a boardroom tool kit to help corporate directors navigate the landscape briefly depicted in this article. (In the interest of disclosure, I am a member of the WEF's Global AI Council.)

Six recurrent themes may prove useful to conducting board-level conversations on AI. Some of these are more appropriate for audit, risk, or technology committees, depending on how oversight responsibilities are delegated. As information technology becomes a bigger element in almost every business strategy, thanks to the convergence of multiple disruptive technologies, a technology committee can expand the time the board can dedicate to overseeing the strategic and risk implications of these innovations. While the oversight of technology might typically be considered part of the audit or risk committee's responsibility, these committees may not have directors with the depth of experience necessary to evaluate and assess these areas. Also, the agendas of many audit and even risk committees tend to be quite full already, and therefore technology considerations may not get sufficient attention.

Take, for example, Banco Bilbao Vizcaya Argentaria (BBVA), a global bank headquartered in Spain. BBVA is recognized as a leader in the digital transformation of financial services. Sunir Kapoor, a director of BBVA and an expert in information technology who I consulted for this article, helped the board create a technology and cybersecurity committee, one of the six committees that provides definition, oversight, and control of the bank's technology strategy and associated cybersecurity risk. The committee is mostly comprised of independent directors and interacts with the bank's chief officers for information technology, information security, and data to ensure there is adequate analysis and critical review of their approaches and proposals. The committee strengthens the monitoring and oversight of the adopted technology approach by possessing competencies and a singular focus on technology and data not usually present on a board. The committee is also helpful in working with peer committees, such as risk or audit and compliance, as well as the full board, to develop their understanding of the implications of technology and cybersecurity.

The agenda for a technology and cybersecurity committee meeting will change throughout the year, ranging from a review of long-term strategy to a more operational short-term view. Regular standing agenda items might include a review of the company's primary information technology elements and key performance indicators. Senior management typically presents on infrastructure, platforms,

AI Priorities

2016

Some 55 percent of surveyed CEOs said AI could add the most value to these business functions:

- Customer service
- Finance
- Human resources
- Information security
- Information technology
- Innovation
- Manufacturing
- Marketing
- Procurement
- Product development
- Risk
- Sales
- Supply chain

2018

The same question resulted in only five business functions selected by at least 55 percent of CEOs surveyed, pointing to a shift from experimentation to more focused investments.

- Customer service
- Information security
- Information technology
- Innovation
- Risk

—Adapted from IBM's *Shifting Toward Enterprise-Grade AI: Resolving Data and Skills Gaps to Realize Value*

applications, and digital transformation plans, as well as regularly reviews security issues: threats, attacks, lessons learned, and mitigations. At times a more in-depth review specific to the company's employment of AI may be carried out to ensure that any unintended consequences of the use of such technology are understood and consistent with the governance framework of the institution.

Let's dive into the six themes that may merit board discussion as AI continues to proliferate.

1. Business Strategy

AI can transform and disrupt many aspects of a company's strategy. Competition, customer interactions, technology, and operations could all benefit from AI, but they will also radically change the workflows, people skills, and the outcomes possible on these functions once AI is embedded into them.

For instance, think of the audit business model today, then imagine it five years from now when AI is embedded into every workflow. Instead of relying on samples to conduct reviews, a sufficiently advanced AI platform can read and review every contract and transaction. Risks decrease as the time needed to manually test for compliance is reduced. The company can instead focus on redesigning workflows knowing they will be fully compliant. Based on the promise of such capabilities, at least one of the leading audit firms has already doubled its technology investment to transform how value is created for their clients, and has changed the skills requirements of its future auditors.

AI's potential has strong implications for a company's strategy. Consider a new threat from a competitor because its use of AI leads to breakthrough performance and, ultimately, market dominance. For example, a major delivery company's packages are no longer delayed by bad weather because its AI works with weather data to automatically reroute shipments around storms. An athletic apparel company can offer customer-designed products in-store that the customer can immediately take home. In both cases, products informed by AI will result in new types of customer interactions to meet evolving demands and expectations.

To fully understand the company's strategic approach to AI, the CEO will typically collaborate with other members of the executive team. Companies distribute the responsibility of technology functions differently, so board members may find discussions relevant with the chief strategy officer, chief information officer, chief security officer, chief data officer, chief technology officer, chief analytics officer, and even chief innovation officer, if any or all of those roles exist within the organization.

Key Questions to Ask

- Is there a map that outlines and quantifies the opportunities and threats of AI? Its existence, and how broad it is, will say much about the company's focus on AI.
- Where, how, and why are competitors using AI, and what happens if they succeed?
- What foundational requirements need to be in place to create trusted AI and to monitor the establishment and progress to support it? These include, but are not limited to, the data platforms the company uses, data strategy, and the approach to overseeing how data is acquired and used. As a rule of thumb, remember that you can't have AI without IA (information architecture).
- Are there business plans in place that articulate processes from pilot phase to scaled deployment in order to prevent unintended consequences of AI? If AI is integrated into critical company workflows without due care and consideration, the end result could be a portfolio peppered with misfires that fail to deliver the desired return on investment.

2. Impact on Culture

Beyond its technology and business strategy implications, AI changes the way people work. Its successful deployment, therefore, depends on having a culture in place that accepts and embraces AI's ability to streamline work. Do employees accept and trust AI? Are there methods in place to avoid negative outcomes from the use of biased information sets or the improper use of data? Is AI being applied in a way that augments current capabilities, and from the employees' perspective results in value for them? Are concepts such as human-centric design, behavioral economics, and choice architecture known and part of the change programs in the company?

A modern change management program is essential to capture the return on investment of AI. Without implementing these human-technology partnerships, AI will not deliver the intended business value because people won't trust it and therefore won't use it. With AI, employees must evolve from being "experience-first" to become "data-first." In other words, they first need to use AI to seek a set of recommendations and then use their experience to make the right decision with the benefit of those AI recommendations.

The challenge of fostering productive relationships between AI and humans goes even deeper, as AI requires people to continue to train it. Here, the problem is that humans fear that AI eventually will take their jobs. But when the dynamics are just right, the results are indeed remarkable. For example, the combination of personal experience and AI helps radiologists to assess potentially cancerous lesions, or insurance professionals to detect fraudulent claims. Curiously, many studies show that performance is best when both people and AI work together, compared to either AI alone or people alone.

In these instances, the practitioners' collaborative training and work with AI have hard and fast results that indicate success, but other applications may not have such a clear line of sight into a positive outcome. It's an important element of board oversight to understand how management is cultivating the right culture that will enable AI to deliver value—especially if companies are significantly increasing their investments in the technology.

Another concern is AI's implications for the skills the company needs and that its employees possess—or lack thereof. In our study *Shifting Toward Enterprise-Grade AI*, 86 percent of some 5,000 global executives said they believe AI will have an impact on the demand for skills in the next five years. Industry productivity improvements will likely trigger shifts in the labor force. Sixty-seven percent of executives expect that advancements in automation and AI technologies will require roles and skills that don't exist today.

Many organizations have already begun to reskill teams, source new talent, and manage existing workforces after beginning to use AI—all with implications for employees. It affects their attitude toward the company and can create an environment of uncertainty. In light of the possibility of uncertainty, boards will need to ensure management is shifting its focus to training and evolving the workforce for AI, as well as to the cultural implications of such a shift. It's not an understatement to say that creating the right strategy for people and culture in the era of AI entails a reexamination of the company's identity: who it wants to be, who it needs to be, and how its employees see themselves in that near future.

Key Questions to Ask

- What gaps exist between current employee skill sets and those that will be needed for employees to thrive in an AI-driven work environment? This should include both the technical and behavioral skills needed to work in an agile workplace.
- What plans are in place to address the workforce and roles most likely to be affected by the growing use of AI in the company?
- How well does the company's culture support a data-first mindset—making decisions systemically based first on data, then on experience—and rapid learning for AI adoption?
- What plans exist to educate the workforce about the ethical use of AI and the data that powers it?

3. Reporting and Controls

AI can significantly enhance the quality and completeness of an organization's financial management and audits. Audit controls are traditionally performed on a "sample basis" because it is not viable to inspect every transaction. AI technology is capable of examining every single transaction and could transform the process. This implies, in turn, that AI could alter the types of controls required and the actual skills needed, as well as the activities performed by people in the internal audit function.

AI's march toward ubiquity also means that corporate controllers will be able to influence where AI is being used across an entire operation, and provide new levers to realize synergies and performance advantages. Once AI is integrated into mission-critical processes, the world of controls needs to expand to address AI and algorithmic risks, bearing in mind that much of this will be based on evolving regulation.

Key Questions to Ask

- Do either the chief financial officer or the auditing firm employ AI and consider implications from AI in their audit controls?
- Is management fully aware of and compliant with existing data-related regulations and potential regulatory frameworks in areas germane to the industry's use of AI?
- To what degree is management aware of all current and planned uses of AI, especially in core products and mission-critical areas? Do the uses of AI create new reporting and control requirements?
- Are AI-specific goals reflected in management's business objectives? Are there plans to keep management up to speed with evolving data regulations?

4. New and Emerging Risks

For all its promise, AI can create new strategic, operational, financial, ethical, legal, reputational, and security risks. As previously

stated, AI algorithms, like humans, may be biased based on the data used to train them or the individuals who trained them. Subsequent algorithm decisions could lead to adverse impact if not properly supervised.

For example, suppose a training data set for an AI loan application included approved loan applications as well as sets of records for some of the age groups that may reflect a low percent of approvals. An AI system trained on this data set would find the correlation between age category and evaluation, and would infer a statistical rule that would automatically handicap that age group for approval. The same bias could occur with a perfectly balanced data set where the system receives training from a few individuals. The biases of those individuals as they apply to approve or reject a claim will also end up in the algorithm, sometimes creating unpredictable AI behaviors.

To make matters more challenging, two additional handicaps get in the way of the application of AI for mission-critical outcomes. First, the outcomes of AI algorithms tend to be difficult to explain since they evolve based on what they learn. It is hard to trust or verify decisions that can't be explained, and as regulation catches up with technology, this will be an area of focus.

Second, mission-critical applications require resilient and secured algorithms that cannot be tampered with. Today, certain AI models are still easy to "fool," and technology to detect tampering is one area of investment for most AI-leading platforms. These so-called adversarial attacks are well documented and imply that the use of unsecured, open-source AI libraries could inadvertently introduce risks that were not obvious in a world more accustomed to using traditional software, which has an auditable set of instructions.

As a result, companies need to fully understand how an AI-enabled application arrived at a decision in order to both maximize AI's potential and mitigate the risks to a company's reputation. Insights from AI can allow a company to craft customer interactions that will greatly enhance the brand, but a single misstep could destroy its reputation in record time. Imagine, for example, the impact to a loan provider's brand if it were discovered that it was routinely denying applications because of bias in its AI training.

Reputation-linked losses are of growing concern. According to one research study, there was a 461 percent increase in reputation-linked losses between 2011 and 2016.

AI has the potential to increase the precision and efficiency of the systems and processes in place to detect fraud. At the same time, it introduces new avenues for security breaches, as well as new requirements to protect and handle properly all data sets within the enterprise.

Another probing question that directors need to add to their cyber-risk oversight regime is whether all data platforms within the enterprise are up to the latest standards of provenance, reliability, confidentiality, and usage monitoring. While these may sound obvious, recent headlines about data breaches suggest that even the easiest questions are not being routinely asked. And with the average cost of a data breach being $4 million, there is no room for complacency.

A board should actively assess management's handling of cyber risk and the resiliency implications of AI as part of its overall enterprise risk management framework—and not at a cursory level, but with a deep understanding of AI's unique characteristics and dependencies.

Key Questions to Ask

- Is there a documented AI risk analysis for operations, finance, reputation, and security?
- Does the cybersecurity plan include specific actions and analyses for rooting out incremental risks potentially caused by AI?
- AI adversarial attacks involve some nearly unidentifiable changes to input data—like a subtle change in an image—that are done in order to cause a significant deviation on algorithm decisions. Are possible adversarial attacks documented for AI use cases in the enterprise? Are mitigation plans in place?
- Considering both the company's products and mission-critical processes where AI is embedded, are there plans to manage liability and reputational risks from AI and data misuse?

5. Ethics and Trust

AI has the potential to significantly increase productivity and improve the quality of life for existing professionals. Used incorrectly, however, AI can inadvertently scale biases. Unleashed indiscriminately, AI can result in significant disruption and is unlikely to generate the benefits expected.

As AI nears omnipresence—embedded in applications, devices, and business processes across organizations of all types—its scope and scale require special attention. This attention must precede the release of a particular AI-enabled product. Ethical AI standards are essential to avoid unintended harmful consequences, and these standards must be articulated as a series of flexible principles that AI systems and their developers, operators, and managers abide by, extending all the way to the foundational data on which AI depends.

The reason why this matters is because, at its essence, AI algorithms work to provide mathematically optimal answers. The moment AI is inserted into decisions that have consequences for

people and society, ethics and morals must become a key factor in those decisions. In many aspects of our society, we don't consider it acceptable to live by mathematically optimal answers—we live in a world of ethical answers, and therefore AI must be taught these ethics.

One of the most widely publicized illustrations of this point: Would any of us buy an autonomous car where the AI was programmed to make a mathematically optimal decision on who would die in the event of an imminent unavoidable accident—the pedestrians, the driver, or the passengers?

As ethics for AI is a rapidly growing area of interest, many thoughtful approaches are being explored by different governing bodies. For example, the *Ethical Guidelines for Trustworthy AI*, created by the European Commission High Level Expert Group on AI, provides a set of emerging principles for companies operating in the European Union.

The WEF has developed a set of five core principles based on input from its global community to guide organizations on the ethical use of AI. Similarly, the Institute of Electrical and Electronics Engineers has created a working group to define a standard for ethical considerations in "Emulated Empathy in Autonomous and Intelligent Systems," as AI systems have matured to be able to detect and simulate human attributes. For instance, an auto manufacturer could choose to embed AI into a vehicle that will detect when a driver is emotionally unstable or too tired and act to prevent an accident by persuading the driver to stop driving. This working group will provide initial views on the ethical considerations in use cases that business leaders need to understand.

Key Questions to Ask

- Are current company conduct guidelines, values, and mission statements being translated into the ethical guidelines for the use of AI?
- Are there plans to provide education about specific ethics guidelines to those working with AI?
- Does management have plans to embed ethics governance and training into AI initiatives?
- If the degree of AI adoption warrants it, should an AI ethics advisory board be established?

6. Governance Implications

The board's governance or audit committees might also consider including AI implications in their scope of oversight given their impact on trust and ethics. One consideration is to include this conversation within the annual evaluation of enterprise-wide ethics and conduct guidelines. As these policies evolve, they will likely encompass new topics, such as algorithmic accountability. Governance committees should also oversee any internal policies and principles for the use of AI created by management.

Boards should expect to become apprised of regulatory issues related to AI. For instance, many countries are enacting laws to protect the privacy of their citizens, and since AI is dependent on data—and in many cases functions best with the widest and deepest data reach—ethical standards for its governance are essential. The development of AI regulations will likely resemble recent governmental oversight of data-related practices. The European Union's General Data Protection Regulation, for example, has caused boards to closely examine how the companies they oversee are complying. As data and AI regulations evolve—sometimes on a country-by-country basis—so will the risk and opportunity for companies operating in those jurisdictions.

Yet even though governments are passing laws, companies, with oversight from their boards, must take a leadership role and proactively govern their use of AI since the technology will change faster than regulations and laws can be enacted.

AI clearly has a key role to play in business—now and well into the future. It holds unprecedented potential for efficiency improvements and new business models, but because its nature is unlike any technology previously embedded into daily life, AI also has the potential for unexpected consequences and disruption on a scale not seen before.

No business should blindly embrace AI, and no responsible board can afford to remain unprepared to understand and oversee its potential.

Jesus B. Mantas is a senior executive at IBM and an independent director of Biogen. The views represented in this article are strictly his own. He can be contacted on LinkedIn by visiting linkedin.com/in/jmantas.

Imagine the impact to a loan provider's brand if it was discovered that it was routinely denying applications because of bias in its AI training.