TruStage™

Artificial Intelligence

Risk overview

As we witness significant growth in the use of artificial intelligence (AI) technologies, it's essential to consider the potential risks and challenges associated with their widespread adoption. While it is unclear how rapidly AI capabilities will progress or how quickly risks will grow, clearly a thorough understanding of AI necessitates a proactive approach to safeguarding your organization.

Harnessing the power of AI

Given the rapid advancements in Artificial Intelligence (AI) and the growing interest in generative AI, powerful language models like OpenAI's GPT family are gaining in acceptance across a diverse segment of industries, including banks and credit unions.

Many credit union executives are looking for ways to embrace the technology to streamline processes, enhance customer experiences, and optimize operations,

There are many ways AI can transform the credit union industry, including:

- Enhanced member engagement
- Fraud detection and prevention (e.g., phishing, bots)
- Credit scoring and risk assessment
- Automated document processing
- Predictive analytics
- Enhanced data security and threat detection
- Incident response
- Compliance automation
- Target-driven marketing
- IT staffing efficiency and costs
- Employee life cycle from recruitment to performance management and employee development

However, the AI transformation does present some significant challenges and risks.

81%

CEOs express concern that the lack of regulations for generative AI within their industry will hinder their organization's success.

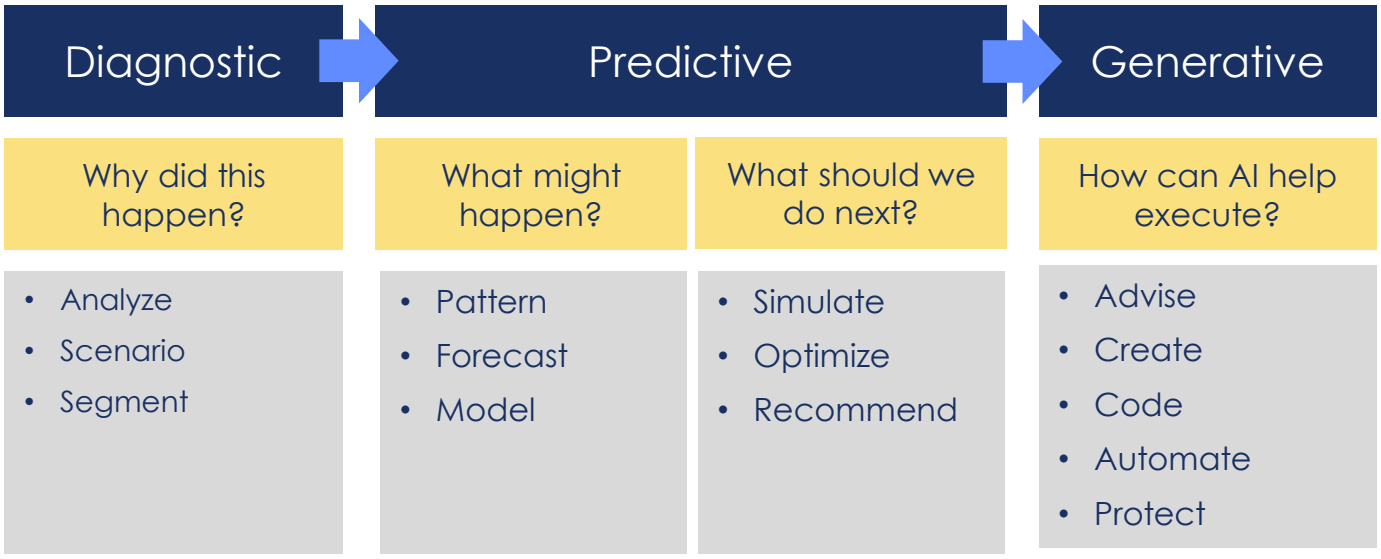
Source: KPMG, 2024

AI risk considerations

- Governance
- Regulation & compliance
- Cybersecurity
- Privacy concerns
- Employment practices (e.g., unsanctioned use)
- Dependence
- Misinformation
- Bias & discrimination
- Job impact
- Fraud (e.g., phishing, deepfakes)

Evolution of AI

AI's evolution is best represented by three phases of significant advancements – diagnostic (introduction of machine learning), predictive (making accurate forecasts) and generative (creating content and offering personalized suggestions).



Source: Work, Workforce, Workers Age of Generative AI Report, Accenture, 2024

Generative AI is a game-changing technology, offering innovative ways to engage users and generate content with deeper insights. It is opening entirely new avenues for improving experiences, delivering new value streams and transforming business models.

With the progress, trends, and growing adoption, there are conflicting views about the risks, benefits and tradeoffs involved with using AI at scale. You must understand the risks and prioritize ethical AI principles, including fairness, transparency, accountability, and understanding.

When planning for and adopting AI and generative AI, it is critical that your organization look at the value and efficiencies it brings. But you also must understand the risks – from those that exist today to those that may bubble up down the road. These are the early days; know that emerging risks will continue to evolve.

The leadership team should set the tone responsibly; however, collaboration and education across your teams will be key to the success of AI.

AI risk considerations

Risk governance & due diligence

Having an AI governance strategy will be vital for your organization. Be sure to include internal and external stakeholders to determine what and how you'll implement AI responsibly.

Consider these questions prior to implementing the use of AI:

- What is the purpose of using AI?
- How will the use of AI be communicated to stakeholders? Is there an AI policy in place?
- What are the ethical considerations for the use of AI?
- What are the legal considerations for the use of AI?
- What are the risks associated with AI use?
- How will the AI policy be monitored and enforced?
- What data privacy and security measures are required for using AI?
- How should AI-driven decisions be validated?
- What processes need to be in place for decision-making and accountability?
- Who will be responsible for the implementation and enforcement of the AI policy?
- Does the policy align with other security and technology policies?

AI models are expected to increase in back-office/IT and finance functions and decrease across employee and HR functions.

Source: KPMG, 2024

Policy implementation

It is recommended to establish a working group or committee comprised of department leaders, executives, and board members to develop a policy, procedures, and oversight that includes:

- Defining your AI risk appetite. Determine how AI may be used at the credit union, to what extent, and which employees are permitted to use it. Consider having employees sign a code of conduct specific to the use of AI, the sharing of information, and the use of company resources.
- Having a method to assess business need/use prior to introducing or allowing AI use.
- Consider restrictions or limitations to how, where, and when AI tools may be downloaded, especially on credit union technology assets. This may trigger software or other technology use security provisions and permissions.
- Expressly prohibit the use, disclosure, and submission/input of restricted information, sensitive information, confidential information, and member/consumer information.
- A requirement to monitor output for the possibility of misinformation: Take proactive measures to combat misinformation in AI-generated content. Internal IT experts should consider creating policies that emphasize the potential risks of misinformation and how AI algorithms work. Policies should further explain the limitations and biases of AI systems.
  - if an AI model is trained on historical data that reflects discriminatory practices from the past, it may perpetuate these biases in current decisions.
  - ensure that algorithms are trained on diverse and representative datasets.
  - how to avoid risk to protected, personal, and confidential information
  - how to avoid violations of copyright/trademark/intellectual property infringements
  - ensuring information and output is free of error, accurate, and compliant.

AI risk considerations

Policy implementation (continued)

- Establish organizational boundaries for AI tools that includes data collection guidelines, privacy and compliance requirements used to query or converse with the AI tool. Does the data provided require member consent?
- Establish messaging for credit union members that explains if and how AI and/or Generative AI can be used in credit union operations.
- Provide credit union members with a mechanism to rate their experience with AI and respond with feedback specific to each service delivery channel.

  For example, in a mobile banking app that uses OpenAI-powered chatbots, one solution is an in-app feedback form that allows users to rate their overall experience with the chatbot and provide specific comments about its performance.
- Require and establish the engagement of human oversight. A team of human experts should review and validate the AI-generated output to catch potential misinformation or harmful content. By actively engaging and collaborating with AI experts, researchers, and providers, you can ensure that your model is continually refined and trained on accurate and unbiased data.
  - Establish an internal review process
  - Document audit findings and any remediation required
  - Set boundaries on the type of content that can be generated
  - Consider diversity and inclusion in the review process to mitigate the potential for bias to be embedded in the output.
- Prohibit or restrict AI access to sensitive member information. For platforms that require access, ensure that robust encryption and controls are in place to ensure member information is protected from threats. Consider regular security audits as a method to detect and resolve vulnerabilities.
- Ensure that any use of AI aligns with member privacy and data protection policies. Consider obtaining explicit consent from members regarding data usage to reinforce a culture of data privacy and accountability. Consider offering clear and easily accessible opt-out options for members that do not wish to engage with AI services.
- Regularly review the effectiveness of the policy and make necessary adjusted based on emerging risks, developments (ex. Regulatory), and information/ feedback from regular review/audits for performance, impact & ethical compliance.
- As compliance and regulations expand, be sure to work with your legal counsel to best understand and adhere to consumer protection risks. Continue to monitor the development of AI, data, and cybersecurity regulation and adjust your usage accordingly.
- Consider providing guidance where to obtain additional information, questions on the policy, and how to report suspected policy violations.
- Expressly provide steps of enforcement and adherence to your policy including action or response to violations of your policy. This may range from suspension of use, discipline, and possible termination.
- Reserve the right to remove, block, suspend, or other prompt act with any AI tools especially if so if it presents a security risk. Such acts may come with or without notice to users.

AI risk considerations

Regulation & compliance

As the regulatory picture around the use of AI continues to evolve, the laws and regulations governing the new technology remain sparse. AI currently lives in an unregulated atmosphere.

There has been some focus in some states and at the federal to create more clear-cut measures to manage the rising sophistication of artificial intelligence. Some courts have also addressed the use of AI. Although legal regulations mean certain AI technologies could eventually be banned.

At this point, it appears that the result may be a complex patchwork of laws and regulations, such as cybersecurity and data privacy rules.

Although the regulatory picture is uncertain, essential compliance obligations can and should be applied.

Many grapple with the ethical and practical issues brought up by the use and misuse of AI technologies.

On a company level, there are many steps you can take when integrating AI into your operations. You can develop processes for monitoring algorithms, compiling high-quality data and explaining the findings of AI algorithms. You can also establish standards to determine acceptable AI technologies.

Core compliance principles such as training, testing, monitoring and auditing are all essential in developing AI policies.

Be sure to include legal counsel, either in-house or external who have the expertise in the relevant areas, because certain existing contracts with data sources and vendors may prohibit the use of some information by AI models.

Cybersecurity

AI has been enhancing cyber security solutions for years. Machine learning models have made network security, antivirus software, and fraud-detection software more potent by finding anomalies must faster than humans.

As AI technologies become increasingly sophisticated, the security risks and potential for misuse also increase. Concerns that hackers and malicious actors can use AI to:

- develop more advanced cyberattacks,
- bypass security measures, and
- exploit vulnerabilities in systems.

Credit unions should strengthen their risk management and cybersecurity practices to counter bad actors' use of AI and consider greater integration of AI solutions into their cybersecurity framework.

Experts suggest that attackers can use generative AI and large language models (LLMs) to scale attacks at an unseen level of speed and complexity. Bad actors can also optimize their ransomware and phishing attack techniques by improving their efforts with generative AI.

Future AI-powered tools may also allow developers with entry-level programming skills to create automated malware, like an advanced malicious bot that can steal data, infect networks, and attack systems with little to no human intervention. Researchers warn that the AI worm represents a new breed of "zero-click malware," as the victim does not have to click on anything to trigger the malicious activity or even propagate it.

Unfortunately, like other cybersecurity issues, an organization that utilizes AI can suffer reputational damage if the technology malfunctions or suffers a data breach. This can also lead to your organization facing fines, civil penalties, and impact existing and potential member relationships.

AI risk considerations

Privacy concerns

AI technologies often collect and analyze large amounts of personal data to customize user experiences or to help train the AI models being used. This raises issues related to data privacy and security.

To mitigate privacy risks, you must maintain strict data protection and safe data handling practices. Since AI tools are often trained on proprietary data, it is important to remember that there is usually no method to retrieve such data once it is out in the open. And leakage of information might violate GDPR, CCPA privacy requirements and intellectual property laws, exposing unaware corporate leaders to additional risk.

Data may not even be considered secure from other users when given to an AI system, as one bug incident that occurred with ChatGPT in 2023.

While there are laws present to protect personal information, there is no explicit federal law that protects individuals from data privacy harm experienced by AI.

Employment practices

Staff in many roles are already experimenting with AI tools to boost productivity, create efficiencies and complete tasks, often without security or IT's knowledge or consent.

43% of working professionals have used AI tools like ChatGPT to complete tasks at work, and 68% hadn't told their bosses.

Source: Fishbowl Insights, 2/1/2023

With AI technologies developing quickly, many organizations have opted for one of two choices when it comes to regulating its use among employees.

One option is issuing a company-wide ban on AI and ensuring technical limitations are in place.

While a total ban seems like a good idea, employees will likely find ways to circumvent it, or they will feel restricted and unable to do their jobs without the tools they believe help them perform tasks more efficiently.

Another option is to educate your staff. As with most cybersecurity practices, a knowledgeable team establishes a stronger first line of defense against incorrect AI use and business risk.

- Ensure all staff is familiar with the AI usage guidelines and corporate policies that have been developed.
- Explicitly tell all employees what constitutes acceptable and unacceptable AI use in the workplace. Explain the potential risks to your organization.
- Understand the data and information being entered into AI tools and make employees aware of the risks.
- Monitor staff to ensure they abide by the rules you have established and understand which roles or processes are allowed to access or use AI tools.

AI risk considerations

Human resources & bias

Generative AI has seamlessly made its way into hiring and talent management. It brings the ability to sort through job applications, picking out candidates through predictive analysis and machine learning. Generative AI is extensively used to automate and streamline different facets of the recruitment process, from applicant sourcing to candidate matching.

However, despite the efficiencies, these AI advancements bring some concern particularly in terms of impartiality and biases inherited from human creators.

In some cases, prejudices may inadvertently make there way into the algorithms. Imagine a scenario where previous hiring managers, driven by biases of gender, age, race or disability, rejected candidates for misguided reasons. The AI tool, could misinterpret these patterns as indicators of incompetence and exclude qualified candidates from underrepresented backgrounds.

For example, Amazon's AI algorithm used all curriculum vitae (CVs) submitted to the company over a ten-year period to learn how to spot the best candidates. Given the low proportion of women working in the company at the time, as in most technology companies, the algorithm quickly spotted male dominance and thought it was a factor in success.

Clearly, relying on AI technology for prescreening can remove the nuances. This can also occur if you set-up your AI tool that 4 years of experience is required. It might screen out candidates who don't have that experience but also those that have other favorable attributes that make up for it.

Just as it's problematic when human recruiters don't check their own biases, not screening the AI tool for biases has its own risks. If recruiters don't use the tools properly and input the right prompts to remove bias, they might increase their recruiting gaps and miss opportunities.

To minimize discrimination and ensure fairness, it is crucial to invest in the development of unbiased algorithms and diverse training data sets within your AI tools.

AI risk considerations

Fraud

Fraudsters are now deploying artificial intelligence to enhance fraudulent scam activity and create deepfakes to increase the effectiveness of their social engineering scams.

Using AI components like ChatGPT, fraudsters can create well-crafted phishing emails. Gone are the days when phishing emails were detected through misspellings and poor use of the English language.

However, users cannot simply prompt ChatGPT to turn out a phishing email because OpenAI has a set of rules that prohibits its use for nefarious purposes. Unfortunately, users are finding ways around these restrictions with the right prompts or inputs. Malicious actors share these specialized prompts, which are often referred to as "jailbreaks", in hacker forums.

Malicious actors have also taken to developing their own LLMs, such as WormGPT and FraudGPT, and advertise them for sale in hacker forums. WormGPT and FraudGPT are specifically designed for malicious activities, such as generating a phishing email, so they are devoid of any restrictions. In fact, WormGPT is frequently referred to as ChatGPT's evil cousin.

Generative AI has also enabled new, more sophisticated forms of digital impersonation. Fraudsters are using generative AI to create digitally altered images, video and audio of people saying and doing things they didn't say or do. They only need a sample image, video and voice sample, which can be found on social media sites, to create a deepfake of the person they intend to impersonate. Deepfakes will be a game-changer for fraudsters in social engineering scams, particularly the business email compromise (BEC) wire scam.

AI risk considerations

Business email compromise (BEC) & fraudulent instruction

Rather than rely on commonly-used phishing emails, fraudsters can use deepfake technology to clone a credit union CEO's voice. Fraudsters only need a sample of the CEO's real voice to create the deepfake. For example, a fraudster could call the CFO - spoofing the CEO's mobile phone number – and, using the CEO's deepfake voice either in a live call or a voice message request a large dollar wire transfer. A request received in this manner will be more convincing than a phishing email since the CFO will recognize the CEO's voice.

AI and deepfake technology will take the BEC scam to new heights.

Another potential fraud scenario could involve a deepfake video. The fraudsters only need a video sample of the person they want to impersonate to create a convincing video with the use of generative AI. Video conference calls between employees of an organization is commonplace. A fraudster could use a deepfake video of a credit union's CEO in a video conference call with the CFO to request a large dollar wire.

Deepfake technology is rapidly evolving to the point where it may be impossible to detect a deepfake voice or video. Nevertheless, there may audio/visual clues suggesting a deepfake.

- Are there long, unnatural pauses in speech? This could be a sign that the fraudster creating the deepfake voice is using a text-to-speech model. It takes time for the fraudster to type the words to replicate.
- Be alert for slurring. The deepfake subject may slur certain words as the technology struggles to create words and phrases that the real subject did not say in their voice sample.
- Does the voice sound natural? Or is the voice monotone lacking in voice inflections that is a part of natural speech.
- Is there background noise? Is the voice crisp and clear or is there background noise like static?

For deepfake videos, pay particular attention to visual clues:

- Is there unnatural blinking by the subject, distorted facial features, or strange lighting or shadows?
- Do the subject's facial expressions, eye movements, head and body movements look natural? Do they match the context of the video?
- Pay attention to lip sync. Do the lip movements and spoken words align?

Deepfake fraud case

Fraudsters clone a company director's voice in a $35 million fraud

This scam involved a two-pronged approach – a deepfake voice and phishing.

A branch manager of a Japanese company in Hong Kong received a call from the company's director – a voice he recognized. However, the director's voice was a deepfake.

The director informed the branch manager that the company was about to acquire a business so the branch manager would need to make $35 million in bank transfers. The branch manager also received phishing emails containing instructions for the transfers appearing to come the director and an attorney allegedly retained to coordinate the acquisition.

Believing the instructions were legitimate – particularly after talking to the director and recognizing this voice, the branch manager made the transfers.

AI risk considerations

Business email compromise (BEC) & fraudulent instruction

As the BEC scam gains traction through the use of deepfakes, its critical for credit unions to revisit their risk mitigation strategies. Credit unions should consider the following:

- Train employees on the BEC scam and how deepfakes are used by fraudsters to increase the effectiveness of the scam.
- Add an external warning flag to incoming emails that originate from outside of the organization.
- Block employees from creating an email rule to send all incoming and outgoing email to the user's trash folder.
- Always – without exception – authenticate the CEO's or other C-suite executive's remote wire transfer request no matter how it is received, whether by email, live voice call, voice message or even video conference call.
- Verify the request face-to-face with the requestor or by calling their extension, mobile phone, or even their home phone.
- Adopt dual approval requirements for large dollar wire requests with one approver being a member of the executive leadership team.

To combat the use of generative AI for fraud, credit unions should consider investing in a solution that can detect deepfakes. Deepfake detection solutions – a growing field of technology solutions - are designed to identify and prevent the spread of manipulated digital content. These solutions are designed to detect modifications and alterations in videos, images, and audio clips that are generated using AI.

Deepfake fraud case

Fraudsters deploy a deepfake video to steal $25 million

A finance employee working in a multinational company's Hong Kong office was duped in a video conference call to make 15 transfers totaling $25 million.

The scam featured deepfake videos of the company's CFO and other employees who were known to the finance employee in a video conference call. The scam also included a phishing email sent to the finance employee that appeared to come from the CFO.

The finance employee initially thought the email was suspicious but was later convinced that the transfer request was legitimate after the video conference call with the CFO and other employees, so the transfer was executed.

Vendor impersonation

Credit union staff may develop a close working relationship with vendors frequently speaking to their vendor contacts over the phone or by video conference calls. Fraudsters could create a deepfake voice or video of the credit union's contact person at the vendor who provides updated banking information for remittances in a live phone call, voice message, or video conference call.

Similar to the risk mitigation strategies for the BEC scam, credit unions should verify changes in banking information for paying invoices received from vendors no matter how the changes are communicated to the credit union including email, live voice call, voice message, or by video conference call. The changes should be verified by performing a callback to the vendor using a reliable phone number.

AI risk considerations

Scams against members

Fraudsters will also likely deploy deepfakes to enhance scams against members, particularly the romance scam. Fraudsters will look to lure more members in the romance scam using deepfake voices and videos to make the scam more convincing.

Scams against elderly members will also become more convincing using deepfake technology.

For example, an elderly member receives a call from someone pretending to be their grandchild (the perpetrator may or may not know the grandchild's name). The "grandchild" indicates they have been arrested, and they urgently need money to make bond. Circumstances may vary or be embellished such as they have unpaid tickets they must pay before being released, or they are calling the grandparent because they don't want their parents to know.

Prepare your employees

No recent - and perhaps no other – technological innovation has transformed so much, including itself, so quickly. Indeed, AI has captured the world's attention.

As your organization integrates AI and generative AI, comprehensive staff training and learning will be vital to ensure your people have market-relevant skills and the capability to collaborate with machines.

After all, humans need to teach the machines (itself a new skill); both people and machines need to get better at their jobs over time to maximize the benefits of AI in this generative stage.

Be sure to prioritize and share with your employees the use cases you have identified and explain how you plan to leverage the AI technology in a positive way.

Looking for additional risk insights?

Risk & Compliance Solutions • **800.637.2676** • riskconsultant@trustage.com