K logix

The AI Wore Tennis Shoes

AI in the Workplace:

A position paper on what organizations need to know and high-level guidance to consider

Table of Contents

Prologue

K logix has drafted this whitepaper for a technical management and executive audience to consider and review as they explore using AI tooling in their organization. Our goal for readers is to takeaway:

- A better understanding of AI tools and how they function
- Clarity around the benefits and risks of these new tools
- Guidance around steps all organizations must take when:
  - Preparing to pursue AI tools
  - Building a program around AI tools
  - Considerations when investing in a specific AI tool

K logix's consulting and research departments, drawing from their work with global businesses and their own extensive research on this fast-moving technology, developed this position paper. While we urge all organizations to practice due care and consider moving cautiously, we fully recognize and appreciate the advancement opportunities that AI tools represent. This position paper will assist organizations with building their policies and programs to reduce risk and safely unlock the benefits of this fantastic new software type.

Introduction

ChatGPT and similar large language model tools are helpful for various tasks such as natural language processing and content generation. However, its increasing use in multiple domains has raised concerns about the potential risks and challenges regarding privacy, security, as well as ethical and legal considerations. ChatGPT counted 100 million active users within just two months of its official launch, according to researchers at UBS[1], making it one of the fastest-growing apps in history. Individuals are not the only users of this application; many businesses are exploring how ChatGPT, and its peers can optimize projects, discover new solutions, and sometimes replace human workers.

With technology transforming rapidly in the workplace, technology leaders are rightfully concerned with the security implications that come with it. Is it safe to use? Can these platforms report on what they are doing with the data they may access? How accurate are the results, and can they be trusted?

This position paper is focused on providing some guidance and direction for organizations looking to establish their position on one of the most significant technological changes of the 21st century. K logix hopes technology leaders will consider our recommendations regarding their business needs and risk tolerance to help their organization chart the path that makes the most sense for them through these swift-moving waters. While, as we will discuss, AI has been involved in technology longer than ChatGPT and similar tools, the recent developments represent a tidal shift that all organizations must be prepared to address appropriately.

What is AI? Why are ChatGPT and Large Language Models Different?

Artificial Intelligence[2] is a concept that has been around almost as long as computers. The idea that a computer program could "think" for itself has been explored, both technically and philosophically by academics on a serious basis for years. Much of that conversation has veered towards the science fiction-sounding concepts of living and thinking machines, such as Blade Runner's Replicants or The Matrix's living machines. However, concepts of AI, such as independently deciding on the best result or mapping thousands of data points quickly to arrive at the most logical answer, have been integrated into the technology we all use every day, so much so that we no longer link it to AI.

AI is utilized in many of the technologies we use every day. We have identified a few categories that contain "AI" technology:

Perception

Technologies that use AI to review images and cross-reference them against databases.
This technology allows for quick identification of figures in a picture, perhaps to be removed, such as in Google's Magic Eraser, or for identifying people removing items from store shelves at

[1] https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/
[2] https://www.techtarget.com/searchenterpriseai/tip/4-main-types-of-AI-explained?Offer=abMeterCharCount_var2

Amazon Go stores. Satellite technology has been using this capability for years to help build out mapping and resource tracking.

Machine Learning

A subset of AI where the program "learns" from each action. Technology such as Alexa or Siri are examples of this as they learn how users speak to them and develop better responses.

Natural Language Processing (NLP)

Turns computer programming to examining the written word. Email filters are a great example of natural language processing that we all enjoy. Users of such platforms such as Gmail and Outlook already benefit from the power of natural language processing (NLP) to block spammy emails as well as make suggestions on the next few words they should type. Natural language processing applies certain logic based on the formulaic language rules to categorize messages as spam or business and generally accepted phrases we may write in an email.

Large Language Models

These are the latest evolution of tools and utilize some of the capabilities of machine learning and natural language processing. Tools like ChatGPT are sophisticated software programs that leverage Large Language Models to quickly search and identify answers to queries. At its core, ChatGPT and tools like it leverage the data in which it has access to mathematically determine the correct response based on the words or prompts supplied. With the large dataset it has been trained on, it can quickly develop the most likely correct response to requests such as "Write me an email canceling my cable service" or "What is the best way to make sushi?" It does this by breaking down the request into participle words and letters, and processing across the data it has been trained on, determining the correct response to the combination.

The above are examples of a concept called "Artificial Narrow Intelligence," which applies software-based thought to processes.

As smart as tools like ChatGPT and the-like may be, they cannot deduce matters they cannot self-validate themselves. For example, you may ask ChatGPT what color the sky is, and based on the information provided, it is highly likely to say blue. Still, it cannot validate by seeing the sky for the color it is and is unaware that it should check itself.

On the other hand, "Artificial General Intelligence" is a higher version of artificial intelligence that would be capable of deducing that the sky is not blue, but grey, due to the cloud cover, and can both validate it and is aware enough to understand that it should double check, perhaps by leveraging an optical data feed or weather reports. We are not at the point where Artificial General Intelligence is available, and once it is, there will be many questions regarding capability, impact on society, and even considerations of personhood for such a program. For this position paper, we are focusing on the concept of Artificial Narrow Intelligence, as represented by the product types listed above, which is rapidly becoming a part of the technology fabric we use daily.

For the scope of this position paper, we are delving into more modern versions of Artificial Narrow Intelligence focused on applying computer-aided thought and reason to solving knowledge-based problems such as analysis, content generation, and code development.

How Are Organizations Using AI?

As listed above, modern AI can assist with many tasks that modern-day businesses engage in. AI can resolve simple and mundane tasks around information. Unlike existing or previous tools, AI-based tooling can perform analytical or development tasks that would require the time and attention of trained professionals and complete it in a fraction of the time a person might take.

Some examples include:

Sales	• Drive Optimization: Ability to target leads to reps who will win them • Conversation Intelligence: Ability to track emotional intonation to determine the success rate of a call
Marketing	• Develop Targeted Ads: Generating narrowly focused marketing content to target niche markets • Campaign Optimization: Make the most out of each campaign and increase profitability • Image Recognition: Leveraging AI models to identify images that align with marketing campaign goals
Information Technology	• Chatbot Improvements: Creating more realistic automatic helpdesk responses • Coding: Ability to write basic to complex code for programs • Code Review: Reviewing large sections of code for errors
Accounting	• Spreadsheet Management: Using AI to build and manage spreadsheets for accounting • Reporting and Analysis: Developing reports based on simple prompts in a short timeframe • Complex Workflow Optimization: Optimizing workflow such as payroll processing to reduce error
Supply Chain	• Inventory Management: Automatically identifying and addressing gaps in supplies • Delivery Optimization: Changing delivery routes in support of forecasted customer needs • Supplier Review: Synthesizing data on suppliers to spot corruption, labor, or cyber security risks

Cybersecurity Benefits of Using AI

There are tremendous cybersecurity benefits to using AI. When appropriately utilized, AI platforms such as ChatGPT may significantly assist organizations looking to increase efficiency and streamline operations. Below are use cases from organizations leveraging widely available tools such as ChatGPT.

Advanced Awareness Training

AI can be utilized to tailor training material to include policies and procedures in funny or serious ways, depending on prompts. A typical ChatGPT use case is for threat actors to create highly sophisticated phishing emails customized to the specific organization they are targeting. However, organizations can use this strategy to improve security awareness training or periodic phishing simulation campaigns. Organizations can provide ChatGPT with parameters to make the phishing email more authentic and test end users' ability to identify and react accordingly.

Policy and Procedure Development

ChatGPT can be a great asset for organizations looking to improve or add to their current policy and procedure libraries. Many organizations choose to, or are required to, follow specific security frameworks or regulatory standards. With ChatGPT, an organization may develop policy and procedure templates specific to their needs. For example, you can ask ChatGPT to create an Access Control Policy specific to NIST 800-171, ISO 27001, CIS, etc. It is important to remember that any templates generated by ChatGPT should be used as a starting point for reasons we will discuss in greater detail.

Going beyond platforms such as ChatGPT, organizations have seen natural language processing integrated into their tools and resources. ChatGPT is an application of natural language processing itself, but many cybersecurity companies are using versions of NLP in their programs to help clients protect themselves. As NLP capabilities advance, there are many ways NLP may be utilized in cybersecurity tools to enhance detection capabilities, reduce manual tasks, and improve the visibility of vulnerabilities.

Examples of emerging capabilities are on the following page.

Emerging Capabilities:

Anomaly detection in cybersecurity tools

Establishes a baseline of what regular traffic looks like at an organization, and alerts on any variants. To do this, it uses an AI algorithm deriving patterns from the data. A challenge with this process is that the AI algorithm might generate many false positives, which can cause alert fatigue among analysts. Applying NLP techniques to anomaly detection could reduce the number of false positives, minimizing the need for manual review. For example, an Intrusion Detection System (IDS) could leverage NLP to analyze logs, which are partially written in natural language, to accurately identify signs of a security incident.

Endpoint detection and response tools (EDR)

Filter and analyze activity on the endpoint to detect and prevent a malicious threat. Its predecessor, anti-virus tools, filtered threats based on a list of known malicious signatures, an example of an AI algorithm that detects patterns data. Legacy anti-virus tools could only detect known threats, lacking the ability to detect any new types of threats. To address that gap, EDR tools today filter on additional datasets, like abnormal traffic patterns, and conduct dynamic malware analysis, when the tool runs suspicious activity in a controlled environment to observe behavior. While EDR tools identify new threats, they may be time consuming, and visibility gaps might persist. NLP further advances the EDR space by implementing fast-filtering methods, such as analyzing printable strings in the malware code, to detecting indications of malware.

Natural Language Processing (NLP)

NLP can also be used to identify vulnerabilities. Developers, for example, can use NLP tools, like ChatGPT, to identify vulnerabilities in their code. Additionally, NLP capabilities augment existing vulnerability management tools by scanning specification documents. Specification documents provide an overview of the software/system/product in question, such as its purpose, a list of features, code design, and use requirements. Scanning these documents provides clues about where vulnerabilities may exist. A vulnerability scanning system that utilizes NLP capabilities could provide a more vivid picture of vulnerabilities on the network.

Risks of Using AI Tools

AI tools have the potential to bring tremendous value to an organization. They automate repetitive tasks and even tackle tasks that previously required human employees' time, such as content development and code review. However, there are many risks to using AI. These risks can be organized using the traditional cybersecurity triad of Confidentiality, Integrity, and Availability or the CIA triad, as it is often referred to.

Confidentiality

Samsung's semiconductor division recently allowed engineers to check source code using ChatGPT. According to Economist Korea[3], three Samsung employees accidentally revealed Samsung's sensitive corporate data, potentially giving OpenAI and its competitors' insights about its technology:

1. An employee discovered a bug in the source code of the download program for the semiconductor plant measurement database and requested ChatGPT's help to resolve the issue.
2. Another employee used ChatGPT to enhance the test sequence for a program that identifies yield and faulty chips.
3. An employee recorded an internal company meeting on their smartphone, transcribed it with a speech recognition application, and utilized ChatGPT to generate meeting minutes.

All three of these individuals are currently under investigation for disciplinary action. This example highlights the security concerns when utilizing ChatGPT, and the importance of implementing robust security protocols to protect sensitive data.

The cloud-hosted nature of tools like ChatGPT compounds the risk outlined above. While the data is hosted on one cloud platform, it is commingled data. Users' requests, prompts, and injects are all hosted in the same database. This data collection is to help the platform "learn," but it is unclear what prevents it from producing and sharing sensitive content when the right prompt is provided. Further, once data is shared, there are no easy ways to remove the data or its impact on the algorithm. Data stewardship and regulations like GDPR with their "The right to be forgotten" requirements are very difficult, if not impossible, to enforce at this time.

Finally, hackers breaching the platform pose a significant risk. ChatGPT, with a robust security infrastructure backed by Microsoft, was breached only a few short months ago. As identified so far, the ChatGPT Breach was limited to only a small subset of customers, including payment data and other account information. Competing startups may need more resources to secure their platforms properly, and this represents a significant third-party risk, primarily if companies exchange sensitive data with the platform. Considering the amount of data hosted on ChatGPT, it is wise to assume they are a prime target for bad actors.

[3] https://mashable.com/article/samsung-chatgpt-leak-details

Limiting the type of information shared with AI tools is common across many organizations, but integrating and using this technology requires some information sharing to achieve the best results. As the plethora of third-party AI platforms increases, the risk of sharing the wrong information with the wrong partner will only increase. Thankfully, some platforms are beginning to add configurations and tooling that limit the release of data, however, this just moves responsibility from the platform to the user to ensure that confidentiality is maintained.

Integrity

AI tools create material with simple prompts, saving time and effort. However, one of the challenges with Artificial Narrow Intelligence, as presented in Large Language Model tools, is the responses to the prompts are both limited to the data the model has been exposed to and has the potential for inaccuracy. With rare exceptions, tools such as ChatGPT cannot respond to prompts focused on data they have not been exposed to, which is why these types of tools, educated on data from before 2022, cannot discuss current events. Further, the process of creating responses, which includes some aspects of randomness, leads to "educated guesses" that may seem accurate but are factually and materially incorrect. Relying on AI tools solely when developing content can lead to disaster as minor inaccurate details that a human being would identify end up going unnoticed.

This becomes more concerning when developing computer programming code, as minor inaccuracies cascade when injected into larger programs. It is partly for that reason that the popular IT website "Stack Overflow"[4] temporarily banned ChatGPT-based answers when responding to questions. Finally, an experiment by CNET[5], the popular technology website, to use AI tools to draft articles led to it issuing 41 corrections to the 74 articles written by an AI program.

Another significant concern associated with content provided by AI programs is the ethical and legal considerations regarding the potential flaws in the databases underpinning the model. The model's training data could potentially contain biases that could be reflected in the generated output. This training data could lead to unfair and discriminatory outcomes, particularly in sensitive domains such as hiring or legal proceedings. There are also legal liability risks associated with utilizing public AI tools. The model can generate content, including text, images, and videos, which can be used for various purposes. However, if the generated content violates copyright laws or infringes on someone's intellectual property rights, the user of ChatGPT could be held legally liable. Image creators are suing an AI-based image tool for improper use of its material for training purposes. The lawsuit alleges that the images the tool is creating are based on existing images, and the lack of acknowledgment represents copyright infringement. The AI tool creators disagrees, citing the differences between its content and the original content, but the matter will likely require the court system to address and identify new standards for recognizing and using creative content. Organizations will need to consider the fidelity of any images developed by AI tools and their potential liability for copyright infringement when considering using them.

Integrity issues can be diminished with careful review by human employees, and there is certainly an opportunity for machine-aware software development or even AI tools that review other AI

[4] https://meta.stackoverflow.com/questions/421831/temporary-policy-generative-ai-e-g-chatgpt-is-banned
[5] https://www.cnet.com/tech/cnet-is-testing-an-ai-engine-heres-what-weve-learned-mistakes-and-all/

tools for accuracy. As organizations build private AI tools focused on their datasets, the risk of copyright infringement diminishes, but accuracy risks remain. The technology has already improved, as seen in the jump in accuracy from ChatGPT 3.5 to ChatGPT 4.0, but the accuracy of unmanaged responses remains a concern.

Availability

AI tools, either public, such as ChatGPT or Google's Bard, or private, such as the Large Language Model Meta AI (LLaMa)-based tools supported by Meta, are very specialized and complex computer programs.

They can require immense computing power (it is estimated that Open AI, the company behind ChatGPT, is spending over $700,000 per day to operate the software[6]), which may increase as data sets increase. With heavy data usage, tools such as ChatGPT may become unavailable or limit the number of prompts it can respond to in a period, even for paying customers.

Even with implementing private AI tools or integrations offered on top of existing platforms, the ability to maintain and ensure success of the operations is limited to the staff available to assist. Specific AI prompts, API connections, and architecture require specialized experience in high demand. Organizations that fail to hire or train the right employees risk losing many of the benefits of AI to poor responses, spotty service, and potential catastrophic system failures as other systems are impacted.

As this technology evolves and develops, availability challenges should diminish. However, it is difficult to achieve the same reliability that other tools, such as cloud hosting and commercial-off-the-shelf software, have performed for the foreseeable future.

What are the External Threats from AI?

Another significant risk associated with the widespread use of AI tools is the risk of cyber criminals using them against organizations. The model's capability to generate text based on the given input can be misused to create convincing phishing emails or social engineering attacks. By utilizing AI tools, threat actors enhance their communication skills, which they typically lack in their phishing emails. AI tools, for example, can create phishing emails that possess coherence, conversational flow, and a striking similarity to genuine messages at no cost, thereby granting even the most basic cybercriminals the ability to heighten the effectiveness of their social engineering attacks. If these generated messages are sent to unsuspecting users, they could be used to steal sensitive information, compromise systems, or infect devices with malware.

Regulation of Data for AI Models

Currently, there is very little regulation protecting data from being consumed by AI platforms. Large language AI such as ChatGPT consumes/scrapes the internet for data, accessing nearly anything on the internet that is open and publicly available. This data collection may be problematic, as there are no guardrails that prevent specific data from entering the platform. Even with the rapid development of AI tools, problems in this area have made themselves apparent.

[6] https://www.businessinsider.com/how-much-chatgpt-costs-openai-to-run-estimate-report-2023-4

An example from across the ocean illustrates the complexities of AI tools and regulation. Clearview AI has been fined 20 million Euros in the EU for using public data (primarily images) to build biometric profiles on users from France[7]. This data could then be sold to private entities, governments, police, and more, which has been viewed as against European data privacy requirements. This fine, which is based on using public images to create private data profiles, opens up new concerns in AI tool usage.

Currently, there is little to no regulation in the United States that would prevent such databases from being created and sold, like with Clearview AI. Data brokers and the data they possess currently have complete control over our data and what can be done with it. Only citizens in certain states, such as California and Virginia, have privacy laws that pave the way for additional AI regulation. This regulatory grey area will most likely be addressed soon as governments, regulators, and the courts step in. However, it creates uncertainty and risk that all organizations looking to use AI tools must consider.

How To Reduce the Risk of Using ChatGPT

While AI has been part of our technology in one form or another for years, the sudden surge in interest in tools like ChatGPT has forced organizations to reckon with this new and improved form of software-based thinking. Should they allow this new technology and risk a potential data breach or other disaster? Should they restrict it and miss out on amazing advances that could make them more efficient and lower their costs? Should they wait and see how this technology unfolds and invest in the winner of what is already a rapidly changing marketplace?

K logix recognizes that this decision is not easy, and it is made more challenging with the lack of vetted, thoughtful advice to consider. K logix, which has a consulting and research team experienced in helping global businesses tackle tough cyber security questions, has developed a multi-step approach to help guide organizations through these confusing times.

The K logix recommendations are broken into three distinct sets of guidelines:

Organizational Preparations: These are the benchmarks K logix recommends all organizations meet before investing heavily in organization-supported AI.

Operating Principles: These are the policies and processes that should be used to guide decision-making with respect to an organization-supported AI.

Investment Considerations: These are points organizations should consider when they are evaluating specific AI solutions.

[7] https://iapp.org/news/a/greek-dpa-imposes-20m-euro-fine-on-clearview-ai-for-unlawful-processing-of-personal-data/

Organizational Preparations

Before deciding to invest and pursue organizational-supported AI, an organization needs to take stock and ensure certain foundational steps have been taken. These have been selected as they represent strong best practices for any size organization in any industry and are directly related to addressing risks around information systems' confidentiality, integrity, and availability.

Encourage Transparency: Organizations encouraging transparency around information sharing will be well-positioned to identify and communicate inaccuracies with AI results. A culture of transparency where mistakes are not hidden and where people feel comfortable asking questions is less likely to see inaccurate or misleading content or code from an AI application being put into the public or production.

Organizations can foster transparency by:

- Having policies or procedures that encourage multiple reviews of products and content
- Promoting a culture where honesty is paramount, and failure is viewed as a reasonable outcome of risk-taking

Establish Data Governance: Data governance is organizing and managing data within an organization. If an organization is to restrict or allow specific types of data to be shared with an AI tool, it first must be able to identify and define the data it is responsible for. An organization that has implemented data governance can answer questions on what data can be shared and why more quickly than an organization that has not.

Organizations can ensure they are on the path to good data governance by performing the following:

- Identifying the different types of data the organization is entrusted with
- Recognizing the additional data regulations the organization must follow based on data types

Secure Software Development: Proper development of software, including code reviews, testing, and other best practices, are important to establish before introducing AI-based tools for code or another programming tool. Secure software development practices should be ingrained in the internal information and application engineering teams and part of any external software product review.

While the hallmarks of Secure Software Development will vary by organization, some best practices to focus on before investing in AI are:

- Implementing security best practices such as limiting admin accounts and applying multi-factor
- Vulnerability reduction before releasing the product (or implementing it internally) is a priority, and code reviews are a key part of product development

Third-Party Cyber Risk Management: An established process for evaluating the cyber risk posed by vendors will assist in properly evaluating the risk of new AI tools and resources. Vetting how their program was constructed, who has access to it, and what security controls are in place to protect any data that may be shared with the application is paramount.

Managing third parties is a multi-faceted effort, but generally, organizations should strive for the following:

- All business units follow procedures that ensure that any third-party that is to receive the data from the organization is properly vetted
- When vetting third parties, special care is given to how data is stored, processed, and secured by the third-party

By being able to ensure that the majority of best practices identified above are in place, technical leaders gain confidence that their organization is positioned to confront and address many of the potential risks posed by AI tools. Prepared organizations advance to the next set of guidelines. Organizations that still require proper preparation should consider restricting organization-supported usage of AI tools until they are at the appropriate maturity or there is a highly pressing business need.

Operating Principles

If an organization has decided to move forward with an organization-supported AI investment, it is important to establish certain principles. These principles clarify how the organization is seeking to use this new technology, what limitations it has put in place to reduce risk, and areas of related focus that must be maintained to ensure success.

- Establish clear guidelines by defining 'Acceptable Use': Organizations should establish clear guidelines for the use of AI tools to ensure that employees understand what is and is not an appropriate use. This can include language, confidentiality, and data security guidelines, and should be informed by the organization's Data Governance Program.
- Provide training and support: Employees should be trained on how to effectively use the AI tool and be given ongoing support to ensure they are using the tool productively, securely and efficiently. This training should cover how to ask effective questions, how to interpret responses, and how to verify information. Investing in external, specialized support needs to be part of the budget so the AI tool receives proper maintenance.
- Monitor usage: Organizations should monitor the usage of AI tools to ensure it is being used appropriately and that employees are not engaging in inappropriate behavior. This can include monitoring for offensive language, inappropriate content, and misuse of data.
- Review API: API connections and AI-related tools must be identified and evaluated before integration into any product, process, or another internal system, both at the initial stage and continually thereafter. Many AI products are being released by smaller startups that may or may not have the same protections as larger platforms such as ChatGPT.
- Connection to the broader internet must be as limited as is reasonable: As these tools connect to the internet, their search can lead to potential injects and data poisoning. A researcher was able to prompt ChatGPT to provide incorrect information about

themselves by posting it to their public bio page[8]. Once prompted, the system provided incorrect data to anyone requesting it. For that reason, data poisoning and potentially malicious code injection are risks that organizations need to consider and limit.

- Protect data privacy: Organizations should take steps to protect the privacy and security of data being used in the AI tool. This may include using secure communication channels, limiting access to sensitive data, and ensuring data is only used for authorized purposes. Additionally, identify which data tiers can be authorized to insert into the AI tool, which must receive permission from data stewards, and which are categorically forbidden to be shared. When using or interacting with an AI system, such as when chatting online with ChatGPT, users must be aware that any information they submit to the system can be processed, retained, and used to give answers to others. If a user shares any personal or confidential information from work with ChatGPT or similar, that information may be stored and potentially shared with or sold to others.

- Establish availability tolerance: AI tools are still ramping up to meet unprecedented demand and are frequently unavailable. Private AI tools may have availability issues due to the custom nature of the software and the different connections to data sets they must maintain. Organizations should establish tolerance levels for interrupted service and ensure business operations utilizing AI tools have built this availability tolerance into their procedures.

- Qualitative review: AI tools are capable of many amazing feats of content production, including drafting entire books based on prompts. As advanced as these tools have become, the nature of their programming inherently creates a risk of inaccuracy in their output, and humans can only review this with the appropriate training. Organizations should mandate that all outputs by AI tools are reviewed by humans that have the correct level of experience (or training) to ensure whatever is shared with the public or put into production is correct and error-free.

Organizations should document these principles prior to investing in an organization-supported AI program or tool. These principles should be shared with all stakeholders and reviewed for organizational fit. Training may be less important when the AI tool is a chatbot focused on very simple information requests, but the right training may be critical when it is used to support the programming for a global SaaS application. Careful alignment of these organizational principles to the organization's culture and goals will also help enhance the likelihood of selecting the appropriate AI tool.

Investment Considerations

When an organization has ensured they are at the right level of maturity and that stakeholders involved have agreed to the appropriate principles, the final step is to select the proper AI tool. This paper cannot identify all the potential tools in this fast-moving marketplace, but K logix has provided key consideration factors when selecting a tool:

[8] https://www.technologyreview.com/2023/04/03/1070893/three-ways-ai-chatbots-are-a-security-disaster/

- Product Fit: There are many different AI tools, and each one touts its special programming, data sets, and unique output capabilities. Before selecting a tool, ask, “What problem am I solving?” if the problem you are solving is not number one on the pain points the tool resolves, it may be worth considering other tools.
- “Public” or “Private”: Similar to the debate around cloud hosting that started about twenty years ago, one of the biggest differences is if the dataset the AI tool is using is publicly available and growing, or private and attached to the organization. There is a risk with both.
    - For the public: Users from all over the world are contributing to the knowledge base, which increases the accuracy of the tool but increases the likelihood that the tool will share answers drawn from your organization’s prompts with other users.
    - For private: These tools rely on your dataset to learn and grow. If your dataset is too small, the effectiveness of these tools may be diminished or take longer to reach the appropriate level of fidelity.
- Integrations: Many AI tools are using other tools to facilitate tasks. These include setting up meeting invites, registering web domains, and interacting with other AI tools to create imagery based on written words. How these connections are established (API, etc.) and what information is shared during these transactions is important to understand and consider.
- Regulation: Organizations should carefully consider the impact of the data regulations they must adhere to before selecting a tool. A low-cost AI tool may be hosted in a country where an organization cannot store data within its borders. Similarly, the features a tool offers, such as recording meetings, may run afoul of state recording rules if proper alerting is not built in. Careful consideration of how a startup technology aligns with regulatory requirements should be part of any process before procurement proceeds.

Is Using AI Safe?

With any new technology, there are always risks and benefits. Due to its complexity and sudden arrival, AI has many stark risks but also many benefits. If an organization is to implement the guardrails identified above and ensure that all stakeholders are aligned in following the guardrails, then a limited exploration of ChatGPT and other AI tools may make sense. The risks have to be weighed against rewards, and similarly, the efficiency needs of the business to utilize this new technology versus their responsibility to be good stewards of data and producers of quality products.

Closing Thoughts

AI tools are a fascinating and fast-moving segment of information technology. Like all information technology, understanding both the benefits and risks is paramount. Further, before an organization considers the pros and cons of a particular tool, it must consider if it is ready to undertake this new technology. Similar to the shift to cloud infrastructure, when poorly managed on-premise networks were moved to the cloud, the vulnerabilities only expanded and, in fact, became magnified. AI tools represent a similar paradigm shift. By considering and implementing the guidance K logix provides, we hope organizations will be able to confidently engage these tools.

About K logix

Founded in 2001 with a sole focus on information security, K logix works with security leaders who want to make the business of cyber security better. Our white-glove approach empowers leaders to advance their security programs and strategically align with the business to reduce risk. Through our Consulting, Research, and Penetration Testing departments we deliver strategy and framework assessments, technology and program advisory, security testing services, and technology resale.

Contributors

Brian Rosmus, Lead Research and Technology Consultant, K logix

Jared Lyons, Senior Research and Technology Consultant, K logix

Leslie Ogu, Senior Cybersecurity Consultant, K logix

Ryan Spelman, Managing Director of Cyber Risk, K logix

Shreeji Patel, Senior Cybersecurity Consultant, K logix

Sydney Solomon, Senior Security Consultant, K logix

Tyler Reid, Senior Security Consultant, K logix

www.klogixsecurity.com