How To Build an AI Security and Compliance Program for 2025

Nick Leghorn
Co-Founder, Secure Start Partners

Secure Start Partners

SecureStartPartners.com
info@securestartpartners.com

Introduction

Every few years, a new technology or a new method of operation comes along that requires the security industry to adjust its approach and implement new processes or techniques to maintain the security of our companies and employees and the privacy of our users. Personally I've watched the shift from on-premesis servers to virtual cloud deployments, the move from monolithic deployments to containerized microservices, the demise of internal corporate networks and the rise of zero-trust concepts and the dominance of SaaS platforms.

The rapid adoption and implementation of Artificial Intelligence (AI) is another such shift. But as an industry we've been here before, and we will undoubtedly have to incorporate new technologies and approaches again in the future as our world continues to change and advance.

The good news is that the generally accepted approach to modern security teams and organizations already gives us all of the functions we need to be able to understand and mitigate the risks posed by this new technology, we just need to properly identify the additional controls and changes that are needed.

As usual, there is no single standard approach from industry experts for what controls to implement and how to make that happen. But most proposed frameworks (NIST AI RMF 1.0), audit criteria (ISO 42001:2023), and regulatory requirements (AI EU Act) propose common requirements and components that can be combined to create an AI Security and Compliance program that reduces risk to the business, maintains regulatory compliance, and minimizes friction on the business operations.

This document provides guidance for how Secure Start Partners has implemented AI Security and Compliance programs at companies in the past, and highlights key areas needing attention for companies wishing to implement similar programs themselves.

Definition of Artificial Intelligence (AI)

The European Union's EU AI Act defines Artificial Intelligence best, in our opinion, and is the definition adopted here:

> An 'AI system' means a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.

This definition includes not only obvious uses of AI such as direct use of OpenAI's models or internally developed systems, but also third party services which use some variety of AI or

machine learning in their functionality. That encompasses the majority of SaaS platforms, operating systems, and other technology in the world right now, which can be overwhelming at first. But the goal here is instead of creating a whole new security program to deal with this new technology, instead to go back to basics and tweak our existing security programs to account for this new technology.

Components of an AI Security and Compliance Program

Typically, you would expect a security program for a specific technology to have prescriptive requirements and common expectations. For example, we expect certain levels of complexity in our passwords, multi-factor authentication for our accounts, patches applied within a certain number of days for our systems… the list goes on. But the rapid advancement and implementation of AI technology means that there really aren't any widely accepted practices yet.

As a result, the majority of the recommendations for how to manage AI systems within a security program rely on having a good risk management process and implementing risk mitigation strategies. Specifically, there are five components of this risk management and mitigation strategy that are commonly highlighted:

- Policy Development
- Vendor Management
- Internal Software Development LifeCycle Management
- External Documentation and Disclosure
- Employee Training and Awareness

The remainder of this white paper will discuss each element in additional detail, with examples attached as appendices for recommendations of what documents and processes to implement.

AI Security and Compliance Policy

The most critical step in building an AI Security and Compliance Program is ensuring alignment within the organization, and that has to flow from the top of the organizational chart.

There's going to be a strong tendency for elements within the organization to rapidly adopt new AI technology, ignoring or avoiding the usual security processes. The marketing and sales teams will immediately gravitate towards automated lead generation and AI assisted marketing messages, which can customize outbound sales communications for each prospect in a way that was never before economically possible. Engineers are going to want to implement AI assistants for their code development, providing suggestions and debugging assistance. Customer support staff are likely to want to generally implement an AI system to trawl through

internal documentation to find previously obscure solutions for customers. Management is going to want to have AI assistants that read their email and summarize information.

All of these use cases are valid. All of these use cases are beneficial and should be encouraged. But all of them represent a risk to the business and need to be appropriately secured.

Without setting the tone from the top with a proper policy, accompanied by the appropriate level of buy-in from management, implementing an AI Security and Compliance program will become a game of whack-a-mole. New AI systems will appear out of the ether, unsupervised and unsecured. Additional pressure solely from the security team will only lead to increased resistance and opposition from employees.

> *The more you tighten your grip, Tarkin, the more star systems will slip through your fingers.*

Before imposing any new requirements, the ground work needs to be laid with management. The company management needs to be on board with the need for implementing this kind of a program and the associated controls, and needs to understand the potential repercussions if it is not followed. Even though this may cause some friction and reduce company velocity, the trade-off is worthwhile.

Some useful items to point out in favor of implementing an AI Security and Compliance Program include:

- **Regulatory Obligations**
  Many countries around the world are rapidly implementing AI safety and security regulations, the most visible of which right now is the EU AI Act which goes into effect in February of 2025. The EU AI Act imposes stricter fines and penalties on non-compliance than the GDPR, and is more broadly applicable thanks to its expansive definitions. Most companies will fall under the regulatory scope of this regulation, and may impose some level of obligation on those companies.

- **Customer Expectations**
  Whether your customers are businesses or private citizens, AI can be a scary concept. There have been a number of high profile data security issues related to AI, and your customers may begin holding you to a higher standard as a result. Getting ahead of their expectations can not only keep you in the running for winning their business – it can also “pull the ladder up behind you” to prevent competition from others who have not made the same investments.

- **Risk Reduction**
  Most obviously, reducing business risk is a driving factor behind implementing this kind of a program. The business needs to make money to continue operation, and it can’t do

that if its trade secrets are compromised, internal documents are made public, and proprietary data is no longer proprietary. This may be the least persuasive of the arguments presented here, but is still a valid pathway to discuss the benefits of an AI Security and Compliance Program.

Once management is on board with the requirements and accountable for ensuring the compliance of their employees, the policy itself should be circulated for approval. A draft version of such a policy is provided as Appendix A in this whitepaper, and encompasses the components we will discuss later in this paper. You should take the time to ensure that it meets your requirements and obligations, review it with your company leadership, and have it formally published alongside the rest of your company policies.

AI Vendor Management

The obvious place to start understanding and mitigating the risk of AI systems is by improving your vendor management process.

We discuss implementing an appropriate vendor management process in another whitepaper, but the basic objectives are worth repeating. We want to build a vendor management system that employees actually want to use rather than one they will actively work to subvert, and so we want to make sure we adhere to the following objectives:

- Rapid and responsive – decisions should take less than 24 hours.
- Easily understood and clear criteria – the decision making criteria should be available within the company and not based on “gut feelings” or other subjective methodologies.
- Never say “no” – even if something is absurdly high risk, provide a pathway for the company to understand and accept that risk. Whether that pathway is ever successful is another matter.

Thanks to the new concepts and technologies used within AI systems, it is worthwhile to include a new set of criteria when performing vendor risk assessments that need to be addressed in addition to the existing criteria. An example vendor risk assessment process is included as Appendix B.

Our recommended approach is to identify what would constitute a “red flag” in your environment and design your questions around those specific red flags. If a proposed vendor or SaaS product doesn’t trip any of those thresholds, it might be something that you would want to approve without additional investigation. But if any of those concerning scenarios are present, it may need to trigger a threat modeling session and further investigation. Ideally the vast majority of vendors should move through the review process without additional investigation beyond reviewing their SOC 2 report or other similar documentation.

For AI systems specifically, the following are the “red flag” scenarios that we specifically recommend you identify and investigate further:

- **Remote AI Systems**
  Some AI systems can be downloaded and operated directly on company assets, and these may represent the least risky version of these tools. Being able to control the data being fed to these systems and where that data goes is an important component of managing their risk. For systems hosted elsewhere, such as SaaS tools or remotely hosted AI systems like OpenAI, this may be a red flag as the company would need to transmit data to these systems for further processing. That relationship, and the security of those companies, would need additional investigation.

- **AI Model Output**
  If the output of the model is something that is read by company employees but needs additional manual action to use, does it really represent a risk? The biggest concern here is the generation of “garbage outputs” or some data that is incorrect or harmful, which may negatively impact downstream systems or customers. The level of review of the output of these tools prior to use should be something that is considered.

- **Data Retention**
  Especially for remote AI systems, there is a question of data retention. Some AI tools openly state that any data provided to them is used for training data for their models, which can be a significant concern for sensitive data like contact information and GDPR related PII. Or, even worse, confidential corporate information. There may be scenarios where sending this data to a third party is acceptable, but only if the risk is understood compared to the sensitivity of the data and the utility of the system.

For emphasis, the effectiveness of this component of the AI Security and Compliance Program is dependent on the prior existence and effectiveness of a vendor management program. If that doesn’t exist yet, or doesn’t operate properly, steps should be taken to first address that foundational issue before implementing these additional components.

Be Proactive: A Recommendation

The best way to secure third party AI systems and prevent “shadow AI” from cropping up (scenarios where employees paid for or signed up for systems without your knowledge) is to be first – provide approved systems and processes that employees can use RIGHT NOW so they don’t have to look for alternatives.

What has worked best in our experience is focusing on three specific use cases and providing pre-approved and procured vendors for each scenario.

Platform and Development AI Use

The first use case is generally for companies that have a product or service that they develop internally, and where they want to implement an AI feature of some sort. It will be inevitable that they want to try a bunch of different AI tools and platforms, and integrate with them quickly.

The best way to help them accomplish their goals securely is:

- **Have Pre-Approved AI Providers**

  This can be as simple as getting an enterprise agreement with OpenAI and allowing for Google Gemini API keys to be generated. Identify what providers you would trust, ensure their settings are configured to be as secure as possible, and when requested provide the credentials.

- **Develop an Approved Set of Testing Data**

  Developers are going to want to try new things, including whatever the latest new shiny tool might be. They are going to want to test it for themselves and see if it might solve their problem or work better as a solution. Don't be a blocker – instead, work with them to develop a set of testing data that they can use without restrictions that doesn't include any proprietary or sensitive data. And when they eventually want to use a new service, ensure it goes through the vendor assessment process.

Software Developer AI Coding Assistants

Developers are increasingly using AI enabled coding assistants to perform their tasks. Some developers are making the jump voluntarily, but other times it is the CTO or head of development that is trying to squeeze more productivity out of the existing pool of talent and requiring them to do so.

These tools can be excellent assistants, but also an excellent mechanism for the unintentional leakage of sensitive company data. Each tool will send your sensitive source code (which, let's be honest, probably includes some API keys and other unfortunate information) to a third party for ingestion and analysis before spitting back recommendations to the developer.

Complicating this problem is that each developer is likely to want to use their assistant of choice, and each one costs a monthly license fee to maintain.

The best solution we have seen is to identify and pre-approve a selection of AI coding assistants (at time of writing, we recommend GitHub Copilot and Cursor) and obtain an enterprise account for them. Simultaneously, work with your finance team to set aside a sufficient quality of budget to allow each developer to have one license for an approved tool for the term of their employment (~$10 per month per developer). Then, work with your IT team to

ensure that each developer can choose a tool of their choice and obtain a license for it, managing the licenses when they transition or select a new tool.

General Employee AI Usage

The hardest scenario may be general employee use of AI. There are a variety of use cases, and supporting each one is difficult. But the best option is to provide some common location where they can go and interact with AI in an approved and monitored manner.

Some options for implementing this include:

- **Define a single approved and secured AI vendor** – Some vendors like OpenAI allow for you to implement an SSO connection and securely enable your workforce to log into their platform and use their AI in a safe and compliant manner.

- **Implement an AI Gateway** – Vendors such as Aim Security offer an AI gateway, where you can allow your workforce to log in via SSO and interact with other AI services securely.

The overall goal here is to provide a friction free way for employees to interact with AI, knowing full well that they will ignore any warnings you give them about restricting data provided to these third party AI systems. At least if you have an agreement with these vendors and can control the platform you can theoretically limit the damage.

Securing Internal AI Development

In our experience, there are generally two paths that a company will take when implementing AI as part of their service or operation: either they outsource it to some third party, or they build it internally.

The first version of that internal AI development, where the actual AI portions are outsourced, seems to be the most popular method at the moment. Companies like Google with their Gemini product and OpenAI with their various AI models have made it incredibly easy for a company to simply integrate with an API and begin taking advantage of pre-built AI systems that are operated by someone else as a service.

In these cases, the process is relatively straightforward. The biggest concerns are handled through the third party and vendor management processes, and what's left is essentially the same as the existing Software Development LifeCycle (SDLC) security controls that we've come to know and love in the recent decades.

The second case, where AI model development and training takes place internally within the company, is significantly more difficult to manage. There is no cookie cutter approach to implementing the right security controls here, so instead the generally accepted best practice is

to implement a risk management program around the use of AI, starting with proper documentation and understanding of the AI system, and ensuring adherence to a set of principles that is approved by the company.

These principles should be documented as part of your AI Security and Compliance Policy, which is in Appendix A. For convenience we'll also document them here in this section. Specifically the recommended principles that your engineers should align their development against are:

- Maintain the confidentiality and security of our data and our customer's data at all times.
- Understand all risks associated with implementing AI systems as part of our business and seek to reduce those risks.
- Use only ethically and legally sourced datasets for the training of our AI models.
- Ensure responses from AI systems are accurate and reliable before providing to customers or being used in our own processes.
- Provide transparent documentation to our customers and users of our AI enabled systems of how those systems work and make decisions.

To ensure appropriate alignment between engineers building AI systems and the stated principles, the typical common components of an SDLC can be used – just with a couple elements tacked on. Specifically, here is where you should be injecting questions and recommendations about AI development and use in your SDLC controls.

- **Design Review**
  Getting security team members involved in the design reviews for proposed systems is an important component of maintaining a secure application, allowing for glaring security issues to be addressed prior to spending time on implementing those features. As part of this design review, the security team should be considering whether the AI principles are being followed and if there are any recommendations for improvements.

- **Threat Modelling**
  An excellent idea for complex or critical systems is to regularly perform threat modelling sessions, where engineers can assist in uncovering any latent security issues or identifying potential sources of risk that may not have been initially considered. Adding AI specific considerations in these sessions would be useful and fruitful.

- **QA / Testing**
  While it might not be feasible to implement robust testing of an AI system as part of the deployment and release process, especially for rapid deployments using CI/CD platforms, it should be something that is considered and tested to some extent. Most QA

testing may only include a “happy path” for proving that their desired output happens. It is just as important to test for potential negative outcomes and how the system might handle “junk” output from the model.

- **Penetration Testing**
  Automated penetration testing has not yet developed to the point where it can try to break an AI system, but that day may be rapidly approaching. In the meantime, hiring a professional manual penetration testing team to test your implementation of an AI system is an excellent idea, specifically focusing on the negative outcomes that might impact your business.

The outputs of these processes and reviews should, ideally, be treated like any other defect or vulnerability and scheduled for remediation according to your documented vulnerability remediation program. If one doesn’t exist, the next best thing would be to document the findings and discuss them with management, ensuring that senior management understands the potential risk of these issues.

It is important to remember that, at the end of the day, your job is not to fix everything. **Your job is to identify and document risks within your organization, communicate them to management, and implement controls that appropriately reduce risk that management wants to reduce**. The business may accept glaring security issues in their systems in exchange for rapid development, and that is a valid course of action for them to take. Provided they understand the potential negative consequences, that is.

Documentation and Disclosure

There are two audiences when it comes to documentation: internal documentation for the business, and external documentation for customers and users of your platform.

Internal Documentation

Generally speaking, the goal of proper internal documentation is to ensure that the business understands the source of data used within their AI platforms, how that data is processed, and what the results of that processing are. This helps form a clear picture for management (and regulators, if necessary) of how the system works, and just like a good network diagram or system architecture diagram it can help maintain institutional knowledge of the systems and clarify their operation in the event something goes wrong.

The basic requirements are outlined in the AI Security and Compliance Policy in Appendix A, but it is worthwhile to dissect those requirements and discuss them in detail here. Appropriate internal documentation of an AI system should include:

1. **The objectives, requirements, and specifications for the AI system.**

   What was the AI system intended to accomplish? How was it designed, and what assumptions were made? This may be especially important to contribute to the defense of future lawsuits and regulatory action, especially if your AI system is later accused of inappropriate behaviors.

2. **All components used within the AI system.**

   Especially in recent years we have started seeing more Software Bill of Materials (SBoM) documentation for our systems, and this is an extension of that expectation. Understanding the components that are used within the system can help track any vulnerabilities that may later arise and help with understanding licensing in addition to the normal third party risk management benefits that come from knowing the components used in your systems.

3. **A testing plan for ensuring the AI system meets design specifications.**

   Nearly every SDLC contains testing steps of some kind, and is a critical component in any SOC 2 or ISO accredited software development program. It's an industry expectation, and any company that doesn't have some kind of testing process is running the risk of platform and system instability. Including the AI systems in this requirement is a natural extension of that risk management item, and should be encouraged.

4. **Appropriate technical documentation for users and consumers of the AI system.**

   Having a user manual for the system is a really good idea. If engineers assume that the system will be used in a way that is inconsistent with its actual use, that's something that should be addressed. Ensuring that all users and consumers of the system know how it works and how to work with it should be a core component of any mature organization.

5. **Any requirements for ongoing monitoring of the correct functioning of the AI system.**

   All software will eventually need to be maintained and updated. Ensuring appropriate documentation to support those operations is critical, and something that may be overlooked especially early in the development process. Including a requirement for clear documentation of maintenance procedures is a good idea.

6. **What logs are required to be maintained for the system and the retention period of those logs.**

   Log management, investigation, and analysis are all core components of a good security program as well as best practices for a mature software development process and SaaS

platform development. Understanding what logs are generated by the system, how they are stored, and what their retention period will be is good for ensuring that any security or stability issues can be properly investigated and resolved.

7. **Data specific documentation, including:**
    a. **Source of the data used in the AI system.**
    b. **Requirements for quality of the data to be used in the AI system.**
    c. **Any preparation requirements necessary prior to use of raw obtained data as part of training the AI system.**

    Modern AI systems are trained on a pool of data. Understanding the source of that data will be important for regulatory reasons as well as potential disaster recovery reasons, as it may become necessary to re-train or recreate those datasets in the future. It is also important from a data retention perspective, as the company needs to understand what customer data is included in these datasets and reconcile that with any contractual obligations for data retention or deletion.

Ideally, all of this information should be contained within a single document or a closely aligned set of documents. Being able to find this information in a hurry can be important, especially in moments of crisis and concern.

External Documentation

For current regulations, one of the biggest requirements imposed on operators of AI systems is the requirement to disclose the nature and intention of the processing of data. Why was this AI system designed? What purpose does it serve? What are my rights, expectations, and obligations for using this system?

There is no current consensus about the proper way to convey this information, but an emerging standard is the AI Transparency Notice. We provide a detailed discussion of that document in a separate whitepaper, but the general idea is that it should be a single document that properly describes the AI system for external customers and data subjects and includes a few specific pieces of information. Specifically, this document should provide:

- Purpose and Objective of the AI System
- Data Sources and Storage
- Model Training Process
- Implementation and Use of AI System
- Identified Risks and Mitigations

- A Mechanism for Reporting Errors and Privacy Concerns

Most companies already have some form of a security whitepaper that they make available for their customers and users, commonly placed on a “trust site” alongside their compliance documentation. Ideally this document should also be posted in that same place.

Employee AI Training and Awareness

Once all that is in place, the final component of an AI Security and Compliance Program is likely the hardest to implement: an employee training and awareness program.

The specifics of what should be included in the curriculum for this kind of a program aren’t necessarily well developed at this time, but broadly speaking it aligns with the common goals of a normal Information Security training and awareness program:

- Ensure employees are aware of, understand, and follow company policies
- Assist employees in identifying potentially risky scenarios and avoiding them
- Help employees protect themselves and the company from potential attacks
- Ensure employees know where to report security issues and how to trigger the incident response process

Once again, the basic components of the security program are unchanged. These same goals are broadly what we rely on for stopping things like phishing and insider threats. What has changed is the content, which needs to be tailored to each specific organization and their scenario.

Conclusion

The technology, threats, and risks that come along with the rapid implementation of AI systems are constantly changing. But by leveraging the standard building blocks of a good security program and sprinkling in a few AI specific changes, we can quickly create an AI Security and Compliance Program that meets the operational, strategic, and regulatory requirements for modern organizations.

Secure Start Partners has assisted other organizations in developing and deploying these kinds of programs in other organizations and can help yours as well. If you would like to learn more please reach out at info@securestartpartners.com

Appendix A: AI Security and Compliance Policy Example

Overview

This AI (Artificial Intelligence) Trust & Safety Program is intended to establish, implement, maintain, and continually improve an artificial intelligence management system (AIMS) within the company. This program defines a set of principles which must be followed by all employees and applied whenever AI systems are used, and a specific set of policies that outline more specific requirements for how to develop, deploy, and interact with AI systems.

Principles

- Maintain the confidentiality and security of our data and our customer's data at all times.
- Understand all risks associated with implementing AI systems as part of our business and seek to reduce those risks.
- Use only ethically and legally sourced datasets for the training of our AI models.
- Ensure responses from AI systems are accurate and reliable before providing to customers or being used in our own processes.
- Provide transparent documentation to our customers and users of our AI enabled systems of how those systems work and make decisions.

Policy

1. An 'AI system' means a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.
2. All third party AI systems must be reviewed and approved prior to any use, and may only be used in a safe and responsible manner.
    a. Each specific use case for third party services must be individually approved by the Security and Legal teams prior to use. Approved use cases will be documented and published for internal reference.

b. Once approved, access to third party AI services are only permitted through approved and secure methods.

c. Third party services must only be given the minimum amount of information and data necessary to achieve the desired output.

d. Output from any third party service must be appropriately and safely handled.

    i. All output from third party services must be appropriately sanitized, escaped, and validated to ensure that it is free from technologically dangerous content at the time it is received and ingested from the service and prior to passing the response to any other company service or system.

    ii. It is the sole responsibility of the company employee (or their manager) implementing the third party service to ensure that all responses from the third party are appropriate and conform to the company's code of conduct and acceptable use policy prior to injecting those responses into any "live" customer communications or other externally visible function.

3. Internally developed AI systems must be designed, implemented, operated, and maintained securely and safely.

    a. Prior to development and implementation, a risk assessment of the proposed AI solution must be completed that addresses:
        i. Any obvious security concerns with the system design.
        ii. The potential consequences for individuals, groups of individuals, or societies that may result from the use of this AI system throughout its lifecycle.

    b. For all AI systems, appropriate documentation must be maintained which must include at a minimum:
        i. The objectives, requirements, and specifications for the AI system.
        ii. All components used within the AI system.
        iii. A testing plan for ensuring the AI system meets design specifications.
        iv. Appropriate technical documentation for users and consumers of the AI system.
        v. Any requirements for ongoing monitoring of the correct functioning of the AI system.
        vi. What logs are required to be maintained for the system and the retention period of those logs.
        vii. Data specific documentation, including:
            1. Source of the data used in the AI system.
            2. Requirements for quality of the data to be used in the AI system.

3. Any preparation requirements necessary prior to use of raw obtained data as part of training the AI system.

4. The company shall maintain appropriate documentation regarding use and development of AI systems for third parties and customers to review, and a mechanism for customers to submit reports about their use of the system. This must include:
    a. A mechanism for users of the AI system to provide feedback about the quality and operation of that AI system, including any adverse impacts.
    b. Reporting of any identified security, reliability, or other operational incidents related to the AI system.
    c. Any other identified and documented obligations for customers, stakeholders, or governing agencies which involve reporting on the use of the AI system.

5. The company shall ensure regular training for employees related to the proper use and development of AI systems designed to elevate their awareness of AI systems to a level appropriate to their role within the company. All employees must be provided this training on no less than an annual basis.

Roles and Responsibilities

Employees

All employees are responsible for reading, understanding, and following the policies documented as part of this program. They are also responsible for questioning the activities of other employees which may not appear to be in compliance with this policy and promptly reporting any suspicions of non-compliance.

Failure to comply with these policies and other company policies or promptly report any instances of non-compliance may result in disciplinary action up to and including termination.

Managers

In addition to the responsibilities defined as part of being an Employee, Managers are also responsible for ensuring that they know and approve every use of AI within their teams and their span of control. Any unauthorized use of AI by an employee may also be attributed to their manager.

Security Team

The Security Team is responsible for the following services and tasks:

- Develop a robust and rapid process for reviewing proposed internal and third party AI systems.

- Provide “paved roads” for rapid adoption and use of approved AI systems.

- Detect and investigate suspected instances of use of unauthorized AI, and report any confirmed findings promptly to management.

References

- ISO/IEC 42001:2023 - AI management systems
- NIST AI Risk Management Framework
- CleAR Framework for AI Transparency
    - https://shorensteincenter.org/wp-content/uploads/2024/05/CleAR_KChmielinski_FINAL.pdf
- AI4SP Transparency Notice
    - https://ai4sp.org/transparency/

Appendix B: Vendor Risk Assessment Process

Third party vendors help us get a lot of stuff done, from hosting our cloud infrastructure to making sure our email signatures are synced up with our job titles. But these vendors can also represent a significant risk to the business.

All 3rd party vendors are required to go through the procurement process prior to use and payment. As part of that process, these vendors must undergo a security review to ensure that their use doesn't open the company up to an unacceptable level of risk. This document describes the questions asked during that review process, expectations for risk mitigation for those vendors, and our security risk management process for 3rd party vendors.

For these questions, **the highest risk value of any answered question defines the risk of the vendor**.

General Questions

Question	Low Risk	Medium Risk	High Risk
Will the vendor have access to any of our data?	No Data, or Specific Selected Documents	Sensitive company data (source code, customer contact details, etc)	Access to customer data, production data
What level of access will the vendor require?	No Access	Access to non-production systems (Salesforce,	Access to production systems

		GSuite, QA Environment, etc)	

Software / Tech Services Vendors

Question	Low Risk	Medium Risk	High Risk
Does the vendor have a current compliance attestation they can provide us?	SOC 2 Type 2	SOC 2 Type 1 ISO 27001	None Provided
Does the vendor integrate directly with any of our platforms or services?	No	Yes, but not any production systems or any infrastructure directly supporting production environments.	Yes, including production systems and supporting infrastructure.
How critical are the vendor's services to our company?	Vendor downtime has no impact to production systems or business critical processes	Vendor downtime can be tolerated for some period of time (days, weeks).	Any vendor downtime has an immediate impact on our ability to operate the business.

Professional Services Vendors

Question	Low Risk	Medium Risk	High Risk
Does the vendor have access to our source code or source code used by our customers?	No	Yes, but only specific repositories where we specifically grant them access.	Yes, to any repository.
Is the vendor able to submit changes to our source code or source code used by our customers?	No	Yes, but only with direct approval and review by a company employee.	Yes, without any restrictions.

Artificial Intelligence / Machine Learning Tools

Question	Low Risk	Medium Risk	High Risk
Where does this AI/ML tool reside? How do we interact with it?	Local model installed on company devices or within our datacenter.	Third party service that requires authentication and/or an API key.	Third party service external to the company, no API keys or authentication.

How are the outputs of these models handled?	All outputs are reviewed by employees, and are not directly integrated into any systems.	Outputs are reviewed by employees and used to inform internal business decisions. No connections to customer facing systems.	Outputs are integrated directly with customer facing systems. OR Outputs are not reviewed prior to being ingested or used elsewhere.
Does this service retain any data from our use of this service?	No.	Only for short periods for troubleshooting. Data is deleted within a set and disclosed period of time.	Yes, data is retained for some purposes. This can include model training or any other use beyond providing the service.

Risk Management Process

The risk that each vendor or service poses to the business is defined by the highest value in the tables above. The process for approving vendors will depend on their risk level, and is described below.

Risk	**Approval**

Low Risk	Approval is automatic with no additional requirements.
Medium Risk	For each vendor being onboarded to the company, manager approval is required to acknowledge and accept the risk posed by that vendor's actions. Managers will be held accountable for the activities of the vendors they onboard and on whose behalf they accept the risk.
High Risk	For each high risk vendor being onboarded, a **Security Assessment** is required. After successfully completing a security assessment and obtaining approval, the vendor can be considered a low risk.