RSAC 365
Virtual Series
Summit

Really a New Mouse Trap? Exploring Risks with Artificial Intelligence

Brett Tucker, PMP, CISSP
Software Engineering Institute, Carnegie Mellon University

This material is based upon work funded and supported by the Department of Defense under Contract No. FA8702-15-D-0002 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center.

DM20-1031

Common Themes Related to Risk
Principles Found in Resilience Management Model

- The risk environment will not contract naturally – the number of risks and complexity will increase
- Organizations must get better at “surviving” with uncertainty
- Knowledge and awareness of risk issues must be distributed throughout an organization
- Traditional tools, techniques, and methods may not work in this environment
- Existing organizational structures may not be agile enough to adapt

The Only Certainty That We Can Count On is Uncertainty.

Defining and Decomposing Risk

More Than Just an Index

Risk calculation is often treated as an index in which (probability x consequence) provides an expected monetary value.

We must analyze risk to find the conditions that drive the consequences to actualize risk.

Cyber Risk Management – Nothing Changes

All Risk Management Processes Have Similar Characteristics

Traditional Risk Management

Regardless of the technology being considered, the risk management process should follow the same fundamental steps and abide by the same principles.

Artificial Intelligence (AI)

Risk Considerations

- Although AI is much like any other new technology that threatens to disrupt markets and industries, there are some contextual considerations for the following conditions

- If the conditions can be addressed, that may reduce or eliminate the risk

- As we discuss each condition, it is important to note that risk interdependency may provide opportunity for efficiency savings

Recall That a ***Risk*** is the Possibility of Suffering Loss and the ***Condition*** is the Term That Describes Vulnerability, an Actor With a Motive, and an Undesirable Outcome

AI Risk Conditions to Consider

- Ill-defined problem statement
- Lack of expertise
- Model-system-data disconnection
- Unrealistic expectations
- Data challenges
- Lack of verifiability

We Will Cover Each in Greater Detail. However, We Must Strive to Identify Others—Especially as Each Organization Brings New Context.

Ill-Defined Problem Statement

Detailed Discussion

- **Risk Condition Description**
- Organizations must constantly adapt to shifting environmental conditions including:
- internal or external threat actors who change their tactics,
- changes in the value of assets
- change in venue.

- These shifts can complicate the problem to be solved and often lead to the risk condition involving an **ill-defined problem statement**.

- **Potential Response(s)**
- Requirements and scope definition must be well understood and documented

- **Scope decomposition** into smaller parts dilutes risk impact.

- **Agile** implementation in development that accommodates frequent changes.

Lack of Expertise

Detailed Discussion

- **Risk Condition Description**
- An organization may be unable to assemble the expertise needed to enable the use of AI-related technology.
- Tasks such as defining the problem, developing the model, collecting data, and constructing systems require skills and expertise that may not be readily available.
- **This risk is not unusual to AI expertise.**

- **Potential Response(s)**
- A proactive talent management strategy must be implemented.
  - Foster educational opportunities
    - Training and education
  - Identify opportunities that provide experience for employees
    - Conferences, projects, research
  - Follow rigorous hiring practices
- Hiring and third party contractors have additional risks to manage.

Unrealistic Expectations

Detailed Discussion

- **Risk Condition Description**
- Some customers may be uninformed or uneducated about AI technology.
- This situation can lead to the **unrealistic expectations** risk condition.
- **Since risk is probabilistic, there is always a chance of error.**

- **Potential Response(s)**
- Customers must understand that the science of AI relies on mathematical modeling that enables automated, risk-based decisions.
- Organizations must review risk appetite and readiness to adopt cutting-edge technologies. (e.g. is the organization willing to withstand AI errors, and if so, how much?)

Model-System-Data Disconnect

Detailed Discussion

Risk Condition Description

- An AI system must be able to sense, collect, and compute the needed data to make a decision.
- At times, a disconnection of **model, system, or data** can result in a system that doesn't meet its requirements.
- This risk condition can trigger a risk incident, such as an AI system that produces poor decisions.

Potential Response(s)

- Organizations must have a proactive and disciplined process for requirements exploration and secure development operations with a flexible and nimble software architecture.
- Agile software development is an example of how developers can build a system through a flexible process that adapts to changing conditions while maintaining model, system, and data alignment.

Data Challenges

Detailed Discussion

- **Risk Condition Description**
- Effective models such as these rely on large volumes of data.
- Models rely on the ***relevance*** **and *accuracy*** of data, the **data challenges** risk condition comes into play.
- • poisoned data
- • biased interpretation of data
- • faulty data collection
- • low volume of data

- **Potential Response(s)**
- The organization must build strategies for how to collect, use, and maintain the data the system uses.
- Must also consider appropriate refresh rates, whether or not to expunge old data, and how well the system accommodates change.
- The organization must accept that:
  - not all data is perfect
  - making data usable consumes significant resources.

Lack of Verifiability

Detailed Discussion

- **Risk Condition Description**
- It is critical that users confirm that the risk-based decisions made by the AI system are appropriate.
- This **verification**, if lacking, may question the potential bias and overall trust the users have in the system or even AI technology.

- **Potential Response(s)**
- Organizations must seek to understand:
- 1. Interpretation of the results may be as important as knowing what results to expect.
- 2. Organizations must work to modify and tune the AI model when errors are identified.
- 3. Organizations must adjust risk appetite to tolerate model corrections without risking that their stakeholders will lose trust in AI.

We are Adopting AI…. What Should We Consider?

Tips on How to Begin

- **Initial steps for adopting AI technology** should include the following:

1. The organization should establish a standardized risk management policy and procedures for implementing that policy.
   - Doing so ensures consistency when adopting new technologies amid the related uncertainties.
- 2. The organization should establish a governance structure where risk-based decisions, such as adopting new technologies, can be made.
  - If a risk governance structure is not yet established, the organization may opt to use other decision-making bodies such as a technology council.
- 3. The organization's risk program must work with executives to understand and communicate the willingness of the organization to take risks so that a reasonable risk appetite bounds the scope of decisions.

References for Consideration

[Cohen 2020]

Cohen, Benjamin. Three Risks in Building Machine Learning Systems [blog post]. *SEI Blog*. May 2020. https://insights.sei.cmu.edu/sei_blog/2020/05/three-risks-in-building-machine-learning-systems.html

[Deloitte 2018]

Deloitte. *AI and Risk Management Innovating with Confidence.* Centre for Regulatory Strategy EMEA. 2018. https://www2.deloitte.com/content/dam/Deloitte/nl/Documents/innovatie/deloitte-nl-innovate-lu-ai-and-risk-management.pdf

HBR 2020]

Harvard Business Review. *The Case for AI Insurance.* The Harvard Business Review. 2020. https://hbr.org/2020/04/the-case-for-ai-insurance

[ISO 2018]

International Organization for Standardization. *ISO 31000:2018 Risk Management Guidelines*. International Organization for Standardization. 2018. https://www.iso.org/standard/65694.html

[SEI 2020]

Software Engineering Institute. *A Risk Management Perspective for AI Engineering.* Software Engineering Institute. 2020. https://resources.sei.cmu.edu/library/asset-view.cfm?assetid=642223

[McKinsey 2019]

McKinsey & Company. *Confronting the Risk of Artificial Intelligence.* McKinsey Quarterly. 2019. https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/confronting-the-risks-of-artificial-intelligence

[SEI 2019]

Software Engineering Institute. *AI Engineering: 11 Foundational Practices*. Software Engineering Institute. 2019. https://resources.sei.cmu.edu/library/asset-view.cfm?assetid=633647

[SEI 2016]

Software Engineering Institute. *CERT Resilience Management Model (CERT-RMM) Version 1.2.* Software Engineering Institute. 2016. https://resources.sei.cmu.edu/library/asset-view.cfm?assetid=508084

[U.S. DoD 2018]

United States Department of Defense. *Summary of the 2018 Department of Defense Artificial Intelligence Strategy: Harnessing AI to Advance Our Security and Prosperity.* United States Department of Defense. 2018. https://media.defense.gov/2019/Feb/12/2002088963/-1/-1/1/SUMMARY-OF-DOD-AI-STRATEGY.PDF

Contact Information

Presenter / Point of Contact

Brett Tucker, PMP, CSSBB, CISSP, MEM, MBA
Technical Manager
Cyber Risk Management
Telephone: +1 412.268.6682
Email: batucker@sei.cmu.edu