RSA LEGAL

FIRST OF ALL, SOME BASICS

i. There is no such thing as AI: It's just a software algorithm utilizing some combination of machine learning, deep learning or neural networks. "We" cannot even define intelligence
   1. *AI: Automated Intent*
   2. *AI: Absolutely Ignorant*

ii. AI is not absolute. It is NOT deterministic. And this can be a problem for non-data scientists. So the best way to explain this to your users and customers is that it is an evaluation of risk - something MAY be bad - and you have tools and people to validate that. AI helps you prioritize this faster.

iii. AI is probabilistic. (Fuzzy, analog, not binary). Errors will occur, no matter what you do.

iv. No one, not even Data Scientists or AI designers, knows how an AI system arrives at its conclusions. But you can see how often the system arrives at the same conclusion given the same inputs and parameters. So that is one way to measure reliability.

v. There is no way to ask an AI, "How did you arrive at that answer?" (See XAI later)

vi. AI is entirely biased (data set) sensitive.

vii. AI-ish/Neural systems have no memory.

viii. If you ask your system a question today and then again tomorrow, there is a good chance you will get a different answer. Is that what you really want? And given that datasets can be real-time it is hard to go back and compare to historical knowledge. And so you still need a human involved in the process.

ix. Errors will compound over time unless you know how to continuously monitor and tune the system.

QUESTIONS TO ASK YOURSELF BEFORE INVESTING IN AI (MACHINE LEARNING) CYBERSECURITY PRODUCTS

I. What problem - exactly - are you trying to solve? **Be specific**: Automation, accuracy, finding things humans cannot, etc. This answer will yield different approaches and measurements of "success" as generally a single AI solution will optimize for one of these areas.

1 *For more information, contact us: Kapil Raina (kapil@aisecurityalliance.org) and Stephen Wu (ssw@svlg.com)*

II. Can you put upper and lower bounds on the problem?
III. What results do you expect?
IV. How will you measure the results (ROI/time)?
V. Is there another, more well-established solution?
VI. Will implementing this solution require special staff in addition to current analysts?
VII. Do you have the capability to train the system in your environment?
VIII. Will you have to increase staff to respond and analyze the outputs?
IX. How much error are you willing to take?
X. How do plan on doing a risk assessment of the system?
XI. Are you willing to utilize a system that cannot explain its findings - expecting you to just trust it? (Would do that with staff?)
XII. Do you employ OODA-loop processes?
XIII. Do you have plans if the system 'goes off the rails'?
XIV. How do you plan on 'explaining' system outputs, as per GDPR?

QUESTIONS TO ASK YOUR VENDOR BEFORE INVESTING IN AI (MACHINE LEARNING) CYBERSECURITY PRODUCTS

I. Can you or your product explain the output of your system? What are the metrics of not only the outputs you will use but also the interim metrics to verify the system is working as planned? For example, the output can be "threat detected" but the interim metrics could be sample files, probability related to non-outputs, etc.,
II. How do you prove your initial dataset is neutral (unbiased)?
III. When asking a vendor, ask how much tunability and details can you see in the ML algorithms (as many times these are black boxes).
IV. How do you know if your AI is not becoming more biased (positive feedback) as it is used? (Discriminatory…)
V. How can you prove the system is giving the 'correct' answers? In other words, how can Trust be validated?
VI. Does the vendor use other customer data for improving accuracy - or how can the dataset be widened for both increased accuracy and decreased bias?
VII. How do you protect against GAN, hostile input or user poisoning?
VIII. Will your system output be culturally and ethically neutral? How can you prove that it will not offend some of my international sites and sensibilities?
IX. Show me why your system is more accurate than competing ones? It may be difficult to get side by side comparisons, so if you can, provide them a historical dataset and note the expected results you will look for (in the case your analysts have found a threat and want to see the accuracy of the vendor system).
X. How does the system adjust for FP, TP, FN, TN fluctuations over time?

BOTTOM LINE

I. AI-ish systems are OK when varying, unmeasurable number of errors are risk tolerable and within your policies: Translation.
II. AI-ish systems, without thoughtful human inclusion, can make some terrible decisions.
III. AI-ish system measures humans and their behavior poorly. (Facial recognition is awful. Judicial and life-affecting applications are lawsuits waiting to happen.
IV. There is no Silver Bullet. We are at the 300baud equivalent of sophistication.
V. There are no set standards (yet) for vendor based AI solutions - so you will have to rely on emerging standards, your own data scientists, and historical knowledge for results

LEGAL CONSIDERATIONS

I. Data breach liability has been an issue for a long time. The California Consumer Privacy Act's statutory damages (claimant's ability to obtain a dollar amount without having to show actual harm) add considerably to risk. Consider AI tools for data breach risk management.
II. Consider the standard of care for managing data. Failing to meet the standard of care may cause contract breaches and negligence claims. Is the use of AI now part of the standard of care?
III. Use auditing against established standards and controls to help build the trust gap. Purchasers of AI solutions can consider adding an audit requirement to a vendor agreement, and vendors may need to meet audit requirements.
IV. Consider compliance issues, such as the IL Biometric Information Privacy Act and AI Video Interview Act.
V. Contracts should capture all aspects of AI services and risks. Purchasers can use them as a tool to hold AI vendors responsible for meeting specific benchmarks.
VI. Cultivate a network of experts that can support AI-related investigations and data breach response efforts.
VII. Developers should not forget analysis of IP issues: The possibility of the patent, trade secret, copyright, and trademark protection and infringement avoidance, as well as contractual protections for data sets as trade secrets or confidential information.
VIII. Implement the use of AI governance tools, such as AI risk assessments, information security risk assessment AI considerations, controls to promote transparency and explainability and minimize bias, and policies and procedures regarding the usage of AI tools and services.