Automated Decision Systems for Cybersecurity and Infrastructure Security

Luanne Burns Chamberlain,
Lauren Eisenberg Davis
The Johns Hopkins University Applied Physics Laboratory
Laurel, MD, USA
Luanne.Chamberlain, Lauren.Davis, @jhuapl.edu

Martin Stanley[1]
Brian R. Gattoni[1]
Cybersecurity and Infrastructure Security Agency
Arlington, VA, USA
Martin.Stanley, Brian.Gattoni, @cisa.dhs.gov

***Abstract*—This paper describes and discusses the impact of using automated decision systems (ADS), or decision automation, on the spectrum from decision support systems (DSS), where a human makes decisions based on analytics generated by the system, to intelligent decision systems based on analytics performed by Artificial Intelligence (AI) and Machine Learning (ML), and further, to fully autonomous intelligent decision systems, where a machine independently makes decisions based on its AI and ML capabilities. Specifically, we examine the use of decision automation in cybersecurity and infrastructure security and present a methodology for determining which decisions should be automated and at which level of autonomy.**

Keywords—Artificial intelligence, decision automation, machine learning

I. Introduction

Automated decision systems (ADS) gather and evaluate information about a situation, determine the need for a decision, identify or develop relevant alternative courses of action, select an action, and then apply the action as a solution.[2] This paper examines the implications of adopting automated decision systems in cyber and infrastructure security domains.

Automation is the use of a machine to do work that might previously have been done by a person. Autonomy is the degree to which a system can function without human intervention; the system operates and adapts to changing circumstances with reduced human participation (semi-autonomous) or without human control (fully autonomous). Decision support systems (DSS) are one type of ADS that interactively aid users in judgment and choice of activities. DSS have been in existence for many years. In the 1960s, researchers began systematically studying the use of computerized quantitative models to assist in decision-making and planning. [1] The field of Artificial Intelligence (AI) has offered many tools for achieving automation and autonomy and has given rise to increasingly sophisticated automated decision-making. ADS are evolving from DSS that assist humans to autonomous decision systems that function without human intervention.

II. Decision-Making

Decision-making is evolving from inherently organic processes performed by humans to one that increasingly incorporates technology advances to improve—in one or more characteristics—the ability to accomplish a goal. The underpinnings of decision-making provide context to understand the implications and constraints of these technical advances.

A. Decision Models

Many researchers draw a distinction between analytical and intuitive decision-making and note that decision automation is best suited to the former. [2] Humans are better at intuitive tasks whereas machines are better in more resource-intensive tasks. [3] Analytical decision-making involves structured methods, information gathering, analysis, reasoning, and logical deliberation. Intuitive decision-making includes imagination, sensitivity, common sense, rumination, instinct, and creativity. [4] Three important decision-making challenges are uncertainty, complexity, and equivocality. [5] Data incompleteness, inaccuracies, and inconsistencies further complicate decision-making processes.

There are three types of decision models: normative, descriptive, and prescriptive.

Normative models are *theoretical*, based on the fields of philosophy and mathematics; they attempt to model how perfectly rational agents should make decisions by determining the highest expected value. Normative decisions are made without consideration of constraints (e.g., Bayesian Networks).

Descriptive models are *subjective*, based on the fields of cognitive psychology and empirical psychological science; they attempt to model how (sometimes irrational) human beings actually make decisions by determining expected utility. Descriptive models consider heuristics, strategies, perceived uncertainty, risk, and gain (e.g., Savage's Theorem [6]).

Prescriptive models are *pragmatic*, based on the field of engineering; they attempt to model how humans should make decisions in practice by considering average outcomes, risks, and subjective probability assessments. They describe feasible

[1] The opinions expressed in this article are the authors' and do not necessarily represent the position of the United States, the Department of Homeland Security, or any other entity.

[2] Derived and customized from the University of Massachusetts definition: "Decision making is the process of making choices by identifying a decision, gathering information, and assessing alternative resolutions."

DOI 10.1109/SPW50608.2020.00048

procedures to make the best possible decisions given uncertainty. This report defines a prescriptive model for ADS.

B. Complexity Model

Ralph Stacey's complexity model offers a method for selecting the appropriate management actions in a complex adaptive system based on the degree of certainty and the level of agreement on the issue in question. [7] We modified the Stacey model to describe decision-making across process unpredictability and the cognitive effort required to understand the process. Processes that are fairly predictable and easy to understand are called *Simple*; simple decision trees or sequences of rules can describe these processes (e.g., password authentication). *Complicated* processes are still fairly predictable but require more effort to understand. Although they include more variables to consider, they can be decomposed into well-defined rules or detailed decision trees (e.g., malicious website blacklisting). *Complex* processes are difficult to predict and may have few or many variables to understand (e.g., determining provenance of a malware sample or classification of a zero-day vulnerability). Anarchy is the term Stacey uses to describe processes that have high unpredictability and are also difficult to understand; we refer to decisions of this type as *Convoluted* [e.g., advanced persistent threats (APTs), advanced software agents, or morphing attacks].

C. Decision-Making Autonomy

Decision-making systems can work in conjunction with humans or autonomously, i.e., independently. As decision-making systems evolve, they will be capable of increasing levels of autonomy.

The autometer in Fig. 1 illustrates the levels of autonomy: Manual, Semi-autonomous (Low), Semi-autonomous (High), and Fully Autonomous. Mature decision-making systems today are primarily Manual (Level 1) or Semi-autonomous (Low) (Level 2). Table I further describes each level of autonomy.

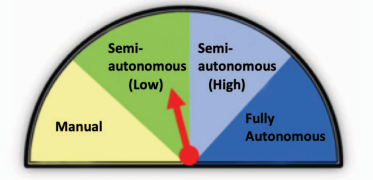

Fig. 1. Autometer

TABLE I. AUTONOMY LEVELS

Level	Autonomy Type	Description	Human-Machine Delegation
1	Manual	User controlled. No automated decision-making occurs; the technology provides automated input to human decision-makers.	Human controlled
2	Semi-autonomous (Low)	User defines the procedures and machine executes them; machine makes decisions based on predefined procedures specified by humans.	Software-assisted human
3	Semi-autonomous (High)	Machine defines the procedures, user approves, and machine executes; interactive decision-making between human user and machine.	Human-assisted software
4	Fully Autonomous	Machine defines the procedures and executes them; machine has the capability to control and repair itself, evaluate the quality of its own work, and adjust its algorithms as appropriate [20]; machine may or may not notify human; machine makes the decision.	Computer controlled

Mature decision-making systems today are primarily Manual (Level 1) or Semi-autonomous (Low) (Level 2). Fig. 2 illustrates the autonomous decision-making overlay on the Stacey Model for Simple, Complicated, and Complex processes as well as for the future notion, fully autonomous systems that are linked with Convoluted processes. Systems with high levels of autonomy could be applied to more complex processes, but they could also be applied to simple problems; Project Maven, a Department of Defense (DoD) AI project [9], is one example—the algorithms are operating with a high degree of autonomy, but the actual task (image classification) is fairly simple and involves only a bounded degree of uncertainty. Likewise, manual decision-making is preferred today in Complex and Convoluted environments because AI is not as capable as humans in those domains.

D. AI and Automated Intelligent Decision-Making

Application of AI to decision-making parallels the maturity of AI capabilities with near-term expectations limited to narrowly focused task areas. According to the National Science and Technology Council (NSTC) Networking and Information Technology Research and Development (NITRD) Subcommittee's national AI Strategic Plan, "virtually all progress has been in 'narrow AI' that performs well on specialized tasks; little progress has been made in 'general AI' that functions well across a variety of cognitive domains." [10] In essence, narrow AI works within a very limited context and cannot perform tasks beyond its field. [11] The Stanford-University–led 100-year study on AI indicates that current AI technologies are highly tailored to particular tasks, where each application would require years of specialized research, and be carefully developed as unique applications. [12] One would not expect the engine that classifies malware to also be used for facial recognition access control. In fact, the algorithms and approaches used to model unique classes of decision automation are quite different. Therefore, great care should be taken when generalizing or reusing narrow AI for other purposes. General AI, which can understand and reason about its environment as a human would, is still on the horizon. [11]

Fig. 3 depicts the relationship between autonomy and AI, as well as other advanced computer science concepts: human and machine teaming, and data science, big data, and ML. AI is advantageous in decision automation as the degree of autonomy increases.

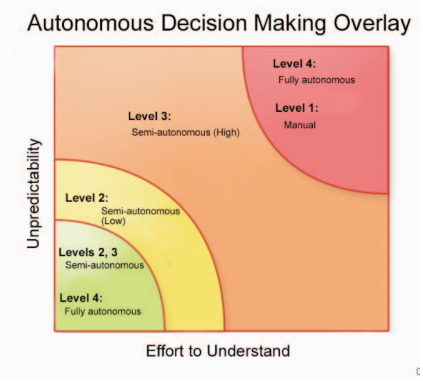

Fig. 2. Modified Stacey Model Autonomous Decision Overlay

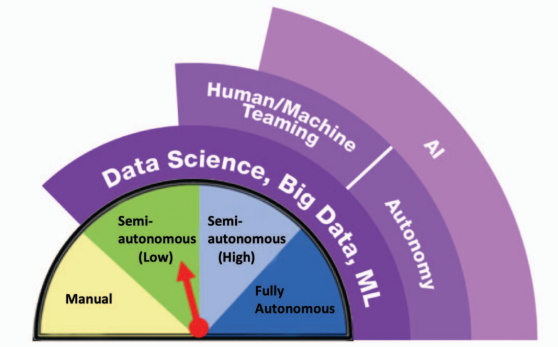

Fig. 3. AI and Autonomy

E. Human-Machine Boundaries

The NSTC report, Preparing for the Future of Artificial Intelligence, states, "Systems that aim to complement human cognitive capabilities are sometimes referred to as intelligence augmentation." [13] In the 1960s, J. C. R. Licklider, an early pioneer of AI, articulated a vision of the relationship between humans and computer intelligence. "Men will set the goals, formulate the hypotheses, determine the criteria, and perform the evaluations. Computing machines will do the routinizable work that must be done to prepare the way for insights and decisions in technical and scientific thinking. The symbiotic partnership will perform intellectual operations much more effectively than man alone can perform them." [14] He believed computers would complement human intelligence. He argued that humans and computers would develop a symbiotic relationship, the strengths of one counterbalancing the limitations of the other.

In theory, fully autonomous system would not need humans in the decision-making loop; instead, humans would provide goal setting, assurance, and verification. Humans would need the skills to oversee what the system is doing, intervene when needed, and maintain the ability to override the machine's actions when necessary. [15] In some contexts, particularly until General AI arrives, humans may continue to provide the algorithms that define the rules for decision-making. In theory, future AI capabilities would be able to handle any cognitive task that humans perform and more. However, the NSTC report states, "In many applications, a human-machine team can be more effective than either one alone, using the strengths of one to compensate for the weaknesses of the other." In reality, many of the decision automation technologies will serve to augment human roles rather than replace them. [10], [13]

Humans will still be required to interpret and explain machine decisions, even if the machine provides a rationale. This skillset is particularly important for cases in which decision automation resulted in the "right" decision based on its parameters, but the parameters were incorrect or the training data did not comprehensively cover the real-world situation.

III. Methodology

Decision automation may not be appropriate for every situation. Care must be taken in identifying the use cases where it would present the best risk/reward.

Before operationalizing an ADS, it is important to evaluate the potential for benefit vs. the potential for risk. This section describes a methodology to systematically assess where to employ decision-making automation and at what level of autonomy.

A. High-Impact, Low-Regret Decisions

The increased adoption of Security Orchestration and Automation Response (SOAR) solutions is driving organizations to evaluate what security response actions to automate; priority should be placed on automating those processes that maximize mission impact and minimize regret if the decision system does not perform optimally. See Fig. 4.

As a Transportation Security Administration example, Baker makes the following statement: "… AI could sharpen security at the landside area of airports. The Evolve Edge system uses a combination of camera, facial recognition and millimeter-wave technologies to scan people walking through a portable security gate. Machine learning techniques are used to automatically analyze data for threats, including explosives and firearms, while ignoring non-dangerous items—for example keys and belt buckles—that users may be carrying." [16] Should AI make the wrong decision and not recognize a firearm being carried by a passenger, the consequences could be grave. This would be an example of a high-impact, high-regret automated decision. Furthermore, the impact of false positives introduces the risk of unacceptable delays and other reputational risks.

B. The Four Dimensions

Determinations for automating decision-making need to be made with consideration to four dimensions: unpredictability, effort to understand, impact, and regret. Fig. 5 shows the benefit vs. risk matrix and provides a decision complexity overlay (from the modified Stacey Complexity model). It also shows the role of humans vs. machines in each quadrant.

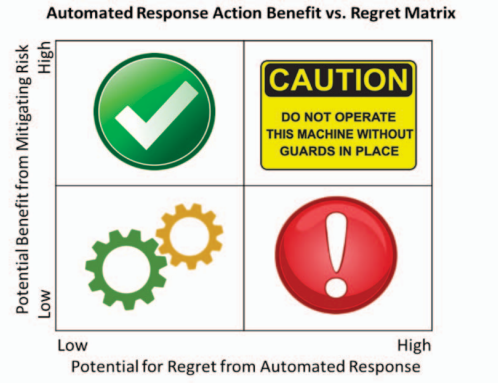

Fig. 4. Regret and Impact

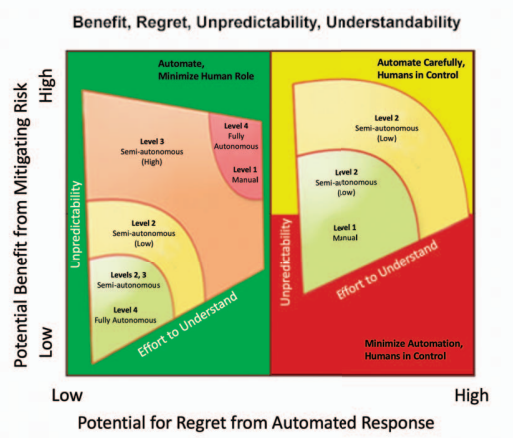

Fig. 5. Stacey Matrix Overlay for Risk-Regret for Automated Decisions

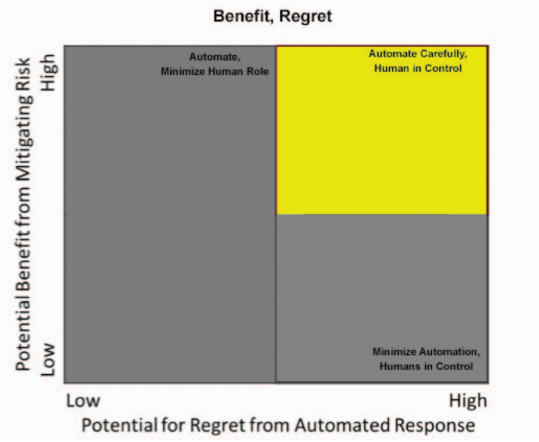

Fig. 6. Plotting Benefit-Regret Quadrant

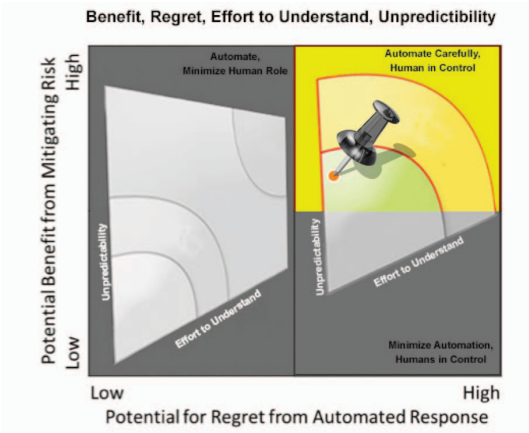

Fig. 7. Plotting Understandability and Unpredictability

The green quadrants represent low risk for potential high benefit. In these quadrants, Level 2 [Semi-autonomous (Low)] systems for Simple and Complicated decisions, Level 3 [Semi-Autonomous (High)] for Complex decision, and Level 4 (Fully Autonomous) systems for Convoluted decisions can be employed. Especially for Simple and Complicated decisions, the role of humans can be minimized in favor or automation.

The yellow and red quadrants are high risk, and the red quadrant indicates low potential benefit minimizing the utility of automation for decision making in those quadrants. Risks that are not confirmed or are of unknown severity, where the associated response action has significant potential to negatively impact the system, are in the lower-right quadrant. The risks of automation are not worth the potential for benefit given the magnitude of uncertainty.

The yellow quadrant, while also high risk, represents high potential for beneficial impact. Humans should still remain in control but can employ DSS to assist them in disciplined decision-making.

IV. Examples

A. *Telephony Denial-of-Service (TDoS)*

For example, an ADS could be used by both attackers and defenders in a TDoS attack. A simple TDoS attack script could be on a webpage or embedded in a smartphone app. The result would dial 911 and, once the call terminates, would loop in an infinite cycle. Additionally, once a determined attack has been thwarted, another attack could be launched automatically by rotating to a different circuit, thereby impeding access to the 911 system for an extended period of time. This situation is compounded when extended to multiple phones placing infinite 911 calls.

Fig. 6 and Fig. 7 illustrate the evaluation of the four dimensions for this example. Note that these types of decisions have the potential for high impact and high regret (yellow quadrant); they are also fairly well understood and are fairly predictable, lending themselves to a Level 2 [Semi-autonomous (Low)] ADS. The evaluation is shown in Table II.

TABLE II. 911 TDoS Attack

Scenario	TDoS bringing down 911 service
Solution	Block illegitimate calls
Decision Automation	Automated blocking
Potential Impact	High
Potential Regret	High
Unpredictability	Medium
Effort to Understand	Medium
Autonomy Level	Semi-autonomous (Low)

The first step (Fig. 6) is to locate the Impact-Regret quadrant by characterizing each as High, Medium, or Low; in this case, these types of decisions have the potential for High Impact and High Regret (yellow quadrant). The next step (Fig. 7) is to evaluate the effort to understand and the unpredictability as High, Medium, or Low; in this case, the decisions are fairly well understood and are fairly predictable, classified as a Level 2 ADS (see the pushpin in Fig. 7).

B. *Incident Triage*

This section presents a cyber incident triage example. Table III addresses a gap in timely incident triage. Each of the four dimensions are evaluated to be high, medium, or low and are then plotted on the corresponding axes. The push pin in Fig. 7 corresponds to the autonomy level of the ADS.

TABLE III. INCIDENT TRIAGE EXAMPLE

Gap #6	Incident triage too slow	
Solution	Automated triage	
Decision Automation	Use semi-autonomous solutions to make decisions about incidents such as classification and whether to pass on to humans or tier 2 analysts.	
Potential Impact	High	
Potential Regret	Medium	
Unpredictability	Medium	
Effort to Understand	Low	
Autonomy Levels	Semi-autonomous (Low)	Semi-autonomous (High)

The potential Impact is High, whereas the potential Regret is Medium; placing the application of ADS in the left, green quadrants of the outer graph indicating a good candidate for decision automation. The Unpredictability is Medium and the Effort to Understand is Low, corresponding to the push pin in the Semi-autonomous (Low) or Semi-autonomous (High) level of autonomy (Fig. 8). Therefore, this is a gap that lends itself to a medium degree of autonomy with potential for high impact and medium regret.

V. IMPACT

The impact of decision automation will be determined by the accuracy, trustworthiness, timeliness and scalability made possible by technology, as the scope of decision-making and related data exceed human capacity.

A. Advantages

1) Timeliness

As decision automation achieves higher levels of autonomy, more timely decisions will be possible. Cyber response actions can be taken without relying on the human analyst to notice underlying characteristics in big data to start the process (e.g., making the decision to escalate incident and perform further analysis, then making the decision to employ mitigation activities for the particular incident).

2) Scalability

As decision-making becomes more automated, the number of decision-making events that can be processed will scale more appropriately to the ever-increasing events that comprise the cybersecurity domain, such as alerts and incidents.

According to Gabby Nizri of the Forbes Technology Council, in addition to custom attacks, IT faces APTs that are increasingly spearheaded by great numbers of automated bots. [17] He asserts that manual response by IT personnel is no match for such intensive, sustained attacks. Not only can humans not keep pace with the sheer volume of incoming threats, they are incapable of making quick and highly impactful decisions to manually address such attacks. Decision automation offers a powerful and effective capability when applied to cybersecurity incident response. To combat the onslaught of incoming threats, organizations must employ an army of equivalent strength and sophistication, especially regarding decisions involving high volume and complex or disparate data.

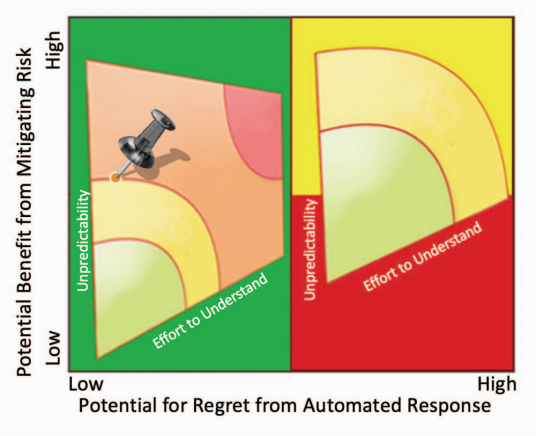

Fig. 8. Four Dimensions: Incident Triage

B. Risks

The perceived advantages of employing ADS do not come without risk. Human decision-makers need to factor those risks and available mitigations into their decisions to deploy ADS. The risk is greater for intelligent decision automation because of reliance on AI and ML. According to David Atkinson of the Institute for Human and Machine Cognition, autonomous systems may perform in ways organizations cannot *a priori* anticipate. [18]

1) Limitations

To account for today's limitations in decision automation technology, users should consider that vendor (or even researcher) claims may be exaggerated. Sufficient preparation must be given to alternative approaches if technology readiness expectations are not realized. One effective mechanism to understand machine performance is the establishment of human performance metrics for the tasks to be automated; however, several points must be noted.

2) Dual Use

Dual use is the ability to apply a technology or concept for both good and malicious purposes. Although many technologies can be subject to dual use, it is of particular urgency when related to autonomy and AI. Decision automation is inherently dual use and must be considered not only for its intended purpose but also for how it could be repurposed inadvertently or by an adversary. Pandya wrote in *Cognitive World* that the dual use of AI causes "enormous security risks to not only individuals and entities across nations: its Government, industries, organizations, and academia … but also the future of humanity." [19] Brundage et al. emphasize that misuse-related concerns must be considered and that harmful applications may not be foreseeable. [20]

3) Emergent Behavior

Decision automation can lead to unintended consequences. Level 3 [Semi-autonomous (High)] and Level 4 (Fully Autonomous) systems frequently exhibit emergent behavior—behavior that is caused by the composition of individual parts into a larger system and cannot be predicted from the properties of the components. The emergent behavior is not found in the individual components.

4) Increased Attack Surface

An ADS inherits the attack surfaces of its non-automated components and introduces new attack surfaces caused by the automated functionality, the AI (cognition and reasoning) code itself, the autonomy code, and the emergent behavior that may result from the autonomous system.

C. Trust

When decision-making is moved from humans to machines, trust must be maintained. Human trust of automated decisions, is based on multiple factors such as policy, ethics, bias, transparency, verification, assurance and explainability. All of these must be considered when employing ADS. "Trustworthiness standards include guidance and requirements for accuracy, explainability, resiliency, safety, reliability, objectivity, and security." [21]

VI. Recommendations

Fig. 9 illustrates seven recommendations for employing decision-automation.

The first recommendation for the near- to mid- term is to consider semi-autonomous systems. The second set of recommendations describes what to evaluate in deciding what to automate. The final set of recommendations is about being prepared in the workforce, increased attack surface, and to examine assurance, trust, ethics and explainability.

VII. Future Work

Having developed a methodology for deciding where to apply decision automation and the recommended level of autonomy, we would like to develop a method for comparing human performance vs. machine performance on decision making tasks.

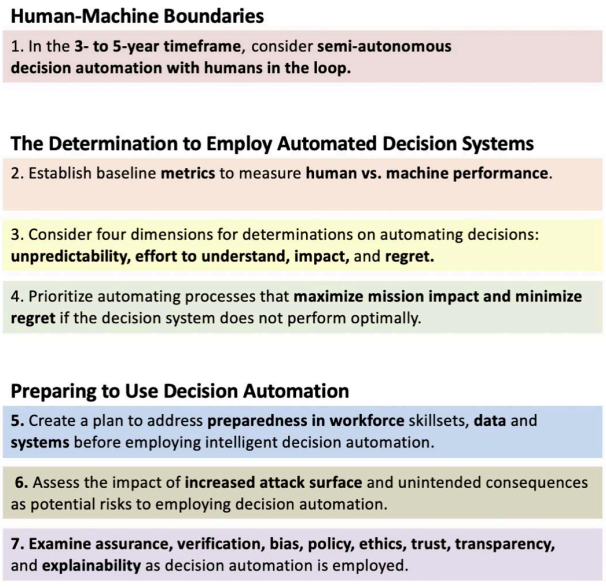

Fig. 9. Recommendations

Cyber-relevant time stresses the importance of communicating information and taking actions in timeframes that are relevant to attack actions; e.g., response at the same rate at which the attacker is attacking. The need for AI and ADS will be driven in large part by the demand for response in cyber-relevant time. Timeliness is not the only measure of performance that is relevant to ADS, however, and approaches for the identification and consideration of these other factors is an area of great importance to the adoption of ADS.

References

[1] D. J. Power, "A Brief History of Decision Support Systems," Version 4.0, 2007 http://DSSResources.COM

[2] JHU/APL, "Continuous Diagnostics and Monitoring (CDM) Phase 3 Request for Information (RFI) Response Analysis," Release 0.0, AOS-16-1013, 2016

[3] S. Mourad and A. Tewfik, Machine Assisted Human Decision Making. *Proc. 2018 IEEE International Conf. on Acoustics, Speech and Signal Processing*, 2018

[4] M. H. Jarrahi, "Artificial Intelligence and the future of work: Human-AI symbiosis in organizational decision making," *ScienceDirect*, 2018

[5] C. W. Choo, "Towards an information model of organizations," *The Canadian Journal of Information Science*, 16(3): 32–63, 1991

[6] Stanford Encyclopedia of Philosophy, Description Decision Theory, 2017

[7] Simple vs. Complicated vs. Complex vs. Chaotic 2008. *Complex Systems*. http://noop.nl/2008/08/simple-vs-complicated-vs-complex-vs-chaotic.html.

[8] M. Ball and V. Callaghan, "Perceptions of Autonomy: A Survey of User Opinions towards Autonomy in Intelligent Environments," *Proc. IEEE International Conf. on Intelligent Environments*, 2011

[9] C. Pellerin, "Project Maven to Deploy Computer Algorithms to War Zone by Year's End," *DoD News*, Defense Media Activity, 2017

[10] NSTC, Networking and Information Technology Research and Development Subcommittee, "The National Artificial Intelligence Research and Development Strategic Plan," 2016

[11] B. Dickson, "What is Narrow, General and Super Artificial Intelligence," *TechTalks*, 2017

[12] P. Stone et al., "Artificial Intelligence and Life in 2030," *One Hundred Year Study on Artificial Intelligence: Report of the 2015–2016 Study Panel*, Stanford University, 2016

[13] Executive Office of the President, National Science and Technology Council Committee on Technology, *Preparing for the Future of Artificial Intelligence*, 2016

[14] J. C. R. Licklider, "Man-Computer Symbiosis," *IRE Transactions on Human Factors in Electronics*, 1960

[15] W. Brantley, "The Data Briefing: Will Artificial Intelligence Tools Replace or Augment Federal Employees?", *Digital.gov*, 2017

[16] J. Baker, "How can AI help speed up airport security?" 2019 https://www.airport-technology.com/features/ai-at-airports-security/

[17] G. Nizri, "Why automation is the key to the future of cyber security," *Network World*, 2016

[18] D. J. Atkinson, "Emerging Cyber-Security Issues of Autonomy and the Psychopathology of Intelligent Machines," *Association for the Advancement of Artificial Intelligence*, 2015

[19] J. Pandya, "The Dual-Use Dilemma of Artificial Intelligence," *Cognitive World*, 2019

[20] M. Brundage et al., "The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation," 2018

[21] NIST, "NIST U.S. Leadership in AI: A Plan for Federal Engagement in Developing Technical Standards and Related Tools," 2019

[22] UMass/Dartmouth, Decision-making process web page, 2019 https://www.umassd.edu/fycm/decision-making/process/195