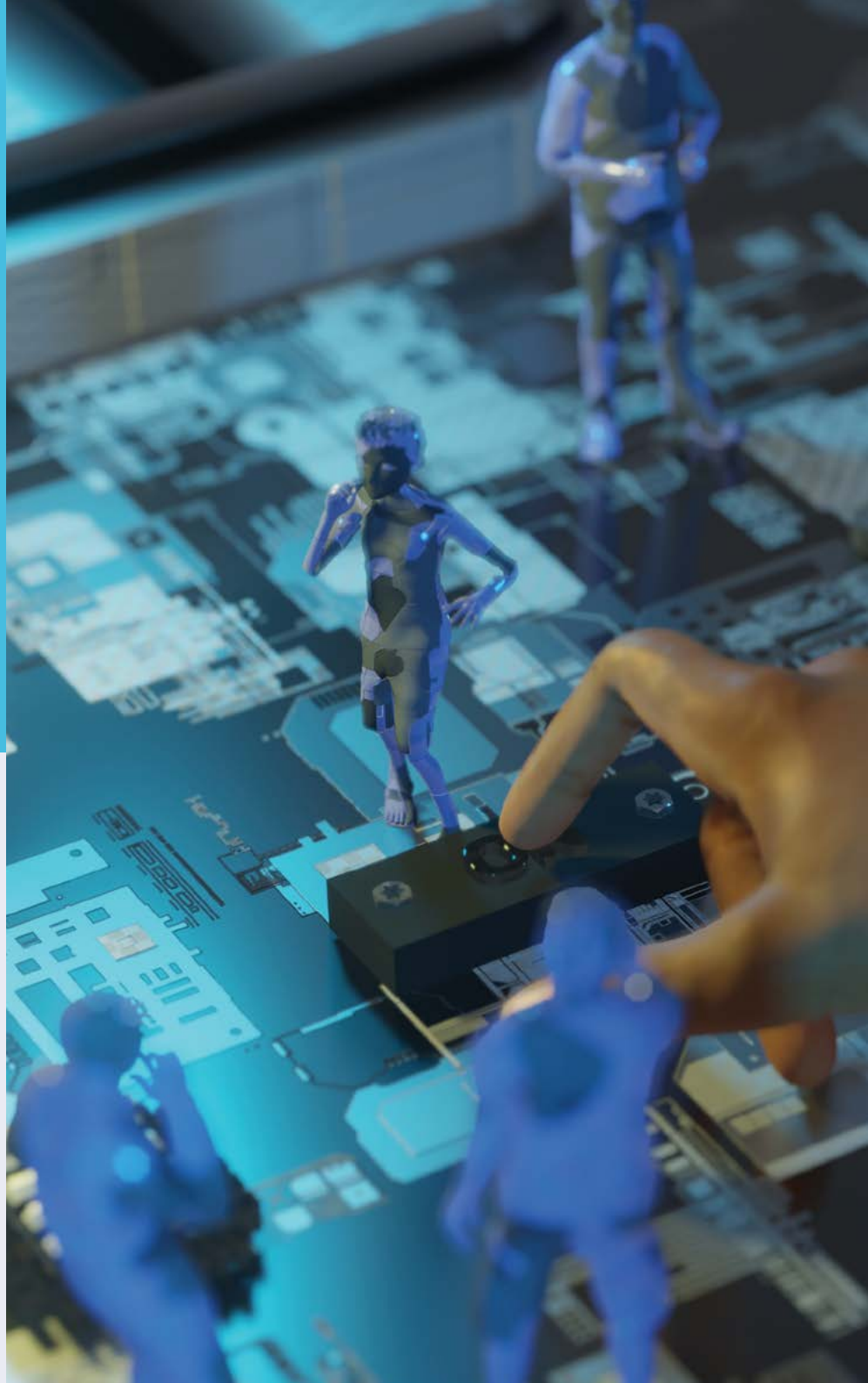
Positionpaper

Checkmate, cyber security?

Karin Bosch, Frank Fransen,
Patrick de Graaf, Dimitri Hehanussa,
Rick van der Kleij, Bert Jan te Paske,
Berry Vetjens and Reinder Wolthuis

An exploration of the potential and limitations of autonomy in cyber security

TNO innovation for life



Checkmate, cyber security?

The malicious use of human-competitive AI for offensive cyber activities poses a clear and present danger to our digital infrastructure. Attackers are already deploying bots for specific tasks and are expected to start launching fully autonomous attacks at scale in the next five years. Like chess grandmaster Garry Kasparov in 1997, cyber security operators may find themselves outsmarted by AI soon. Deep Blue is here. We must prepare for the end of the age of AI-competitive humans.

This paper explores the potential and limitations of autonomous cyber security as a countermeasure against autonomous attacks. Going beyond the idea of automating individual security steps, the security chain in its entirety will be in scope: we consider the organization as a whole and its direct environment as well. We start with exploring the larger trends and the concept of autonomy. Then we introduce Athena, a comprehensive concept for autonomous cyber security: our answer to cyber security's 'Deep Blue moment'. And what technical and societal, legal, and ethical considerations, both in terms of challenges and limitations arise behind the concept of autonomous security? Finally, we touch on alternative courses for a secure digital society in the future.

Context and problem statement

Security has always been an arms race between defenders and attackers, and this is also true for cyber security. Due to the ever-growing complexity of IT systems and the increasing and global connectivity between IT systems, ever cyber security defenders. have an increasingly tougher challenge. With the arrival of autonomous attacks, this challenge may realistically become too complex in the next few years. Especially when the seemingly structural, (global) shortage of skilled cyber security professionals is considered.

Cyber security professionals today need a deep understanding of systems that are developed, managed, and maintained through long and complex supply chains. Networks and systems move to adaptive software defined architectures, and cloud or other third-party service providers are woven into business processes. Boundaries of what constitutes "your" organisation, and "your" IT are getting vaguer. Despite this practically unmanageable complexity, dependencies on IT still increase and the failure of IT systems has more and more real life and physical impact.

Like chess players, cyber security operators face a need for speed, a need for scale (handling ever larger amounts of dynamic data from a growing number of sensors and external sources) and a need for accuracy (ignore false positives, choose the right responsive action out of many 'Courses of Action', or CoA).

Automation of individual cyber security steps and actions has already helped to improve our security posture: "Software is becoming better at detecting abnormalities and vulnerabilities in the system that it is protecting. The sheer amount of data to be monitored for such detection is already far beyond human abilities, and the speed at which things happen limits what humans can do."[1] Automation and application of machine learning (ML) and artificial intelligence (AI) are helpful instruments to support human operators.

From Deep Blue to AlphaZero

In 2017, 20 years after Deep Blue, AI-based chess engine AlphaZero became the strongest chess playing entity on earth. Applying deep learning technology, it reached a level far surpassing that of any other chess software. Let alone us humans, who by then had grown used to being humiliated by the machine.

Automating chess is relatively easy, as the rules of the game are simple and fixed, and therefore easy to model.

In comparison, the cyber security battlefield is not only complex, but also highly dynamic. Cyber criminals are constantly devising new attack vectors. And even more importantly, the systems and networks we defend are rapidly evolving from locally hosted hardware and applications to virtualized, cloud based (micro-) services. Cyber security resembles a game of chess where both the board and the rules change continuously while we play, and where opponents do not play by the rules.

1 Liiovja, Rain and others, Autonomous Cyber Capabilities under International Law, NATO CCD CoE, 2019, p. 22

What is autonomy?

In general, autonomy means self-regulation or self-governance – the ability of a system to establish its own rules of conduct and then to follow them. However, definitions can differ substantially between (scientific) disciplines[2], which can lead to misunderstandings in societal discussions. Especially in technical fields, the meaning of autonomy remains much less clear but generally refers to something less elaborate than autonomy in its philosophical sense.[3] Elements of autonomy include: having sensory input, being able to make informed decisions, drawing on a knowledge base, rules and/or algorithms, goals, ability to act, adaptability to external changes.

2 Tim Smithers, 'Autonomy in Robots and Other Agents' (1997) 34 Brain & Cognition 88; Willem F.G. Haselager, 'Robotics, Philosophy and the Problems of Autonomy' (2005) 13 Pragmatics & Cognition 515

3 Liiovja, 2019, p. ..

Autonomic computing

Amongst others, IBM explored self-managing autonomic (IT) systems, defining the following four types of property:[4]

- Self-configuration: Automatic configuration of components;
- Self-healing: Automatic discovery, and correction of faults;
- Self-optimization: Automatic monitoring and control of resources to ensure the optimal functioning with respect to the defined requirements;
- Self-protection: Proactive identification and protection from arbitrary attacks.

In autonomic systems the **"human operator takes on a new role: instead of controlling the system directly, he/she defines general policies and rules that guide the self-management process."** Another important concept introduced by IMB's autonomic computing vision, was the MAPE-K control loop to manage the adaption of the autonomic systems. The MAPE-K references four steps: Monitor, Analyse, Plan, and Execute, that are supported by common Knowledge component.

4 Poslad, Stefan (2009). Autonomous systems and Artificial Life, In: Ubiquitous Computing Smart Devices, Smart Environments and Smart Interaction. Wiley. pp. 317–341.

In this paper, we consider autonomous security operations as the ability of a system to perform certain operational cyber security tasks without requiring real-time interaction with a human operator.[5] This implies that how the autonomous (security) system performs, is the result of the design and development of that system and the sensory input it receives and not from human decisions. Yet, meaningful human control of the system is a key requirement of autonomous security systems.

5 Likewise Liiovja, 2019

Introducing Athena: autonomous cyber security done right

Imagine, a ransomware attack is attempted at your organization, a large energy company. Attackers have obtained some user credentials through phishing and found their way into your network, in search of interesting information to steal and vulnerable systems to lock with their malware. If they succeed, not only will you see your systems going down, but also your clients might lose their energy supply. Fortunately, "Athena"[6], your autonomous security system, detects the initial malicious access almost immediately, correlating vast quantities of data of several sensors in your IT-networks. It responds with lightning speed, autonomously considering several options. It chooses the best option, based on its knowledge of the IT infrastructure, but also on the impact of each of those options on the internal processes of your organisation and suppliers, and the impact on your clients. Without your own end-users or clients ever knowing, the attackers are kicked out and systems restored and patched and the account with which credentials were stolen is reset. Business goes on like usual for you, the company and your clients.

The Athena system encompasses a complex, wide-ranged autonomous cyber security "system of systems". It covers a wide set of cyber security tasks, which can take place simultaneously. We view it as a socio-technical system (of systems), where effective human-machine interaction and control is a critical design factor.

To describe Athena and its functions we use a well-known and widely used cyber security framework that was developed by NIST (see figure below). The NIST Cyber security Framework provides a structured approach to help determine and address highest priority risks to your business. It identifies 5 functions that form a cycle: Identify, Protect, Detect, Respond, and Recover. Athena's tasks include all elements of this cycle, but on top of that Athena needs to perform two extra 'meta' functions to be effective. These overarching tasks are 'Orientation' and 'Learning and sharing'. In the following paragraphs we provide detail on how Athena can deal with these functions.

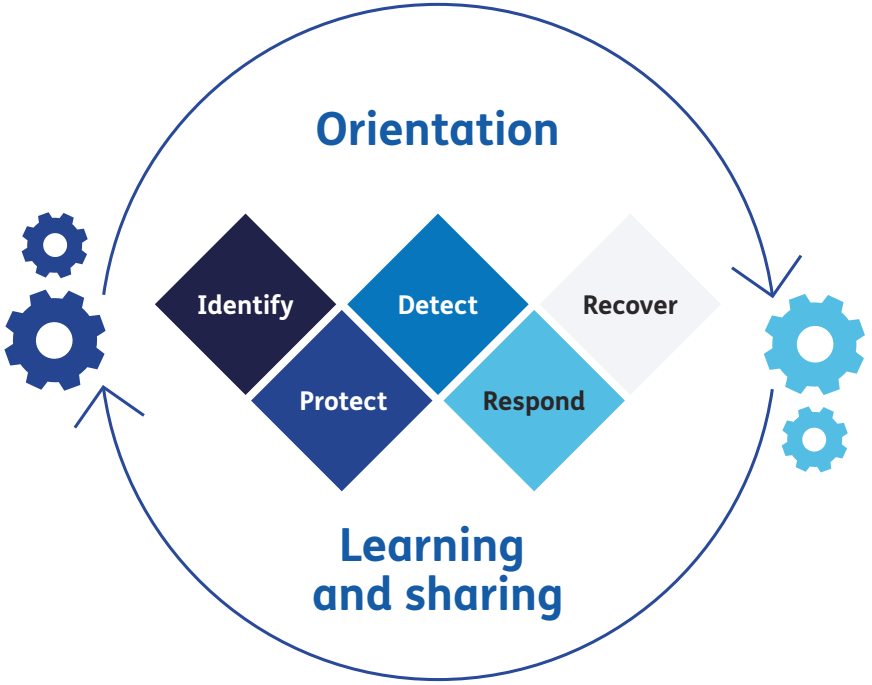


Figure: NIST framework functions, extended with Athena meta-functions

6 Goddess of wisdom, and warfare. So why not for smart cyber warfare? Coincidentally, Athena is also the name of the autonomous decision support system featuring in the techno-thriller "Ghost Fleet: A Novel of the Next World War" by P. W. Singer and August Cole (2015).

Orientation (meta function)

Like for AlphaZero, orientation is a first task, set for Athena itself. It should learn what the to be protected IT environment looks like and how it interacts with business processes (for business impact estimation). Another task for Athena is learning how the organisation is linked to external IT environments (for instance of clients or service providers) and what those environments look like. This understanding may be incomplete, but must be updated frequently. By design it has to deal with multivendor environments.

Orientation also encompasses other data science/AI preparations, like identifying and implementing (incl. learning) the right algorithms and rules for every stage. Furthermore, the relevant networks and systems ought to be equipped with sensors to feed Athena.
Also, continuous training of human operators on Athena and their roles is important, to ultimately maintain control.

Identify and protect

These two steps concern technical and organisational measures prior to cyberattacks or incidents. An autonomous cyber security system can play a role in for instance:

- Automated discovery and patching of known and unknown software vulnerabilities. This should be a continuous and comprehensive activity of Athena across the whole network, as part of a proactive security stance.
- Identity & access management to support the processes that manage the digital identities of users, thus controlling access to critical information or systems. At a minimum, Athena should be "aware" about access & identity policies and monitor access control decision.
- Acquiring and analysing cyber threat intelligence;
- Establish and maintain an asset inventory and network topology. Moreover, establish and maintain a software inventory for each of the assets. An important development here is definition of Software Bill Of Materials (SBOM). For Athena establishing these inventories is very important, as you can't defend what you don't know. Athena should thus continuously collect and
- Vulnerability management and secure configuration management are also cyber security operations that Athena should play a role in. Automatically scanning systems for known vulnerabilities and verifying that systems are deployed in a secure "hardened" configuration is best practice cyber security hygiene.
- Next to establishing an asset and software inventory, Athena should establish an inventory of cyber security controls and functions that are directly enforceable during an cyberattack or to mitigate a vulnerability when no path is available. These so called cyber security actuators are very important for Athena, as these enable Athena to respond. Athena should be able to evaluate if she has sufficient security actuators to respond effectively and efficiently, and with limited to no business disruptions.
- Applying strong security measures for Athena itself.

Detect

Monitoring and detection is arguably the area in cyber security where machine learning has been applied the most and longest. Detection algorithms can already automatically identify non-compliance to "business rules" and anomalies in network, end-point device, and user behaviour. Usually this is applied in more mature organisations. Monitoring here means that Athena listens to its sensors throughout the network and gathers data from external inputs in the form of (amongst others) cyber threat intelligence and relevant changes in (interfaces with) external IT environments.
Detection of suspicious events should trigger a chain of actions by Athena, so that is a crucial element. The sooner an intrusion is detected, the better chances to minimize impact.
False positive signals are an important challenge here. Reducing the number of false positives to near zero seems a precondition to allow for a human-out-of-the-loop system. Detection by Athena has to be actionable, meaning that it is possible to determine a response. Detection of anomalies is not actionable. The anomalies has to be automatically further assessed to confirm that it a cyberattack.

Respond and recovery

- The incident response and recover phases are currently highly depended on human specialists of the so called Computer Security Incident Response

Teams (CSIRT) and IT system operators. Interesting research and development work has been done on the automation of response against cyberattacks or incidents.[7] Automation of Response and Recovery can for instance effectively be applied to create synergy by following the four phases of the MAPE-K loop. The previous sections addressed the monitoring phase of the MAPE-K loop. But note that also during the Respond and Recovery, monitoring continues and is added to better track the actions of the attacker. Respond and Recover is a cyclic process that typically includes many cyclic actions to contain the attacker, eradicate installed malware and remove (root cause) vulnerabilities, and re-build system and install back-ups.
In the Analysis phase, automatic analyse of the security events will need to take place by collecting additional data, assessing the threat, determine what assets are compromised, and determine the potential business impact.

- In the Plan phase, automatic generation of possible responses, so called courses of action (CoAs), to respond to the ongoing attack. The CoAs are assessed on effectiveness and business trade-off. As described above, there are different phases in incident response: containment, eradication and recovery. The first CoAs are thus focussed on collecting more data on what the attacker's actions in order to come up with a good containment strategy. When the system decided on the best containment strategy it execute the CoAs (Execute phase of MAPE-K). The next CoAs will subsequently focus on eradication and recovery, or self-healing. It is important to keep monitoring for malicious activities to determine if the containment strategy was successful.

A specific type of response is the so-called (automated) active cyber defence. This can be described as 'direct defensive action taken to destroy, nullify or reduce the effectiveness of cyber threats against friendly forces and assets'.[8] Active cyber defence may include deploying measures outside one's own networks to counter malicious cyber activity.[9] In general however, due to legal and technical constraints, it seems that application of autonomy in this sense is limited to one's own networks, for most organisations.

Cyber Grand Challenge

In 2016 DARPA issued the Cyber Grand Challenge, to explore technology for autonomous active cyber defence and offence. The technology focus was on automated discovery of vulnerabilities in software, as well as exploiting and patching these vulnerabilities. The winner of that competition, the Mayhem Cyber Reasoning System[10], was a prototype designed to operate in a simplified operating system specifically developed for the Cyber Grand Challenge and demonstrated interesting capabilities for autonomous passive and active cyber defence features. Besides the US, also other nations are investigating autonomous cyber defence capabilities.[11]

10 Thanassis Avgerinos and others, 'The Mayhem Cyber Reasoning System' (2018) 16 **IEEE Security & Privacy** 52.
11 See for instance Japan, https://www.cybersecurityintelligence.com/blog/japans-new-ai-based-cyber-defence-system--4907.html

7 See for instance the SOCCRATES project (https://www.soccrates.eu) and the Automated Security Operations Program (ASOP), www.tno.nl/nl/digitaal/digitale-innovatie/trusted-ict/automated-security-digitale-economie/

8 Dorothy E Denning, 'Framework and Principles for Active Cyber Defence' (2014) 40 **Computers & Security** 108, 109.

9 Robert S Dewar, 'The "Triptych of Cyber Security": A Classification of Active Cyber Defence', **2014 6thInternational Conference on Cyber Conflict (CyCon 2014)** (IEEE 2014)

Learning & sharing (meta function)

To keep up with developments in its environment, Athena should firstly learn by itself from external events, its own actions and their impact on IT and business processes, and thus reconfigure and optimise itself. Most probably this continuous process is (semi) supervised by human operators. Athena should secondly be able to learn from other (trusted) Athenas[12] and share its (her?) own lessons. Not just in terms of sharing cyber threat intelligence, like Indicators of Compromise (IoC), Techniques, Tactics and Procedures (TTPs), and Campaigns, but also how CoA's worked out and how they can be improved.

Athena should be able to operate its various functions in parallel, as one part of the network might require preventive actions, whereas another part might be under attack and in need of response. This also supports the cyber security and IT departments to be both proactive and reactive. This has implications for performance and scale.

12 Or other sources.

Bringing Athena to life

Although envisioning a technically functional Athena is not far-fetched at all in this timeframe of high-performance AI and widespread automation, the road to bringing it to life is not straightforward at all. In this paragraph we introduce a number of technical and ethical/legal/societal considerations that are key to reach the required level of autonomous cyber security.

Technical considerations

How can we design, build, and run a system like Athena? Nowadays many elements are available in the disciplines of cyber security, AI/Data science, psychology, and others. Some of those elements may be on the shelf, others still in the phase of early research.

To start, to function properly, the hunger for data of Athena is incredible (need for scale!). Modelling of IT infra and business processes into machine-readable data is essential. So are sufficient well-placed sensors, for actual view on situation. Applying Sun Tzu, how well does Athena know itself, know its environment and know its enemies?

- Know yourself: have an up to date, accurate and machine-readable model of one's IT (/OT/IoT) infrastructure. Software defined architectures, hyper-connectivity and cloud services raise the question where 'know yourself' blends into 'know your environment'. What are the boundaries of one's own IT? This is not just a technical issue, it also impacts freedom to act autonomously and independently.
- Know yourself: automatically modelling and reasoning on business impact should be accurate and up to date. And here also, is the focus on your own organisation still sensible? Given the fluid boundaries between organisations (and their IT), growing dependence on external IT services, supply chain effects & attacks etc.
- Know your environment: building on the previous two points, Athena should also be able to model Cloud and other 3rd party services to a certain (to be determined) extent. This is new and no doubt complex. Furthermore, it raises questions about availability of data and reliability and trustworthiness of available data.
- Know your environment: it gets even more complex when Athena must reason about the organization's responsibilities/commitments towards external parties.. How do you do that?
- Know your enemy: in this area the cyber industry has a better track record. Acquiring and sharing cyber threat intelligence is fully in development. The main challenge here seems to be selecting the right sources for Athena in the vast supply of free and commercial threat intel feeds.
- And lastly, know and reason over actions, being potential response options and their impacts (CoA's). Such actions are dynamic, as Athena should continuously develop its arsenal of CoA's.

Building a system like this would require a long learning curve, that continues under operational conditions. The challenge is to get representative data in sufficient quantity and quality to learn/train algorithms initially. An interesting feature would be the ability to reliably learn from other autonomous security systems.

All those models and data on actual states needs to be fed into a database structure of some sort and be accessible real-time (need for speed!). Technical performance with this large amount of dynamic data quickly becomes an issue to deal with.

Performance might well be impacted by the need to protect the autonomous system itself against cyberattacks and counter AI attacks.[13] A system like Athena is the ultimate target for advanced attackers. Protecting ML learning processes against data poisoning, system integrity, communications (authentication, encryption) etc. poses a challenge. Who guards the guardian?

Connectivity for communication of Athena within and outside the organization is key as well. The whole system is dependent on sensory input, third party input and the ability to set out actions in the network. A design consideration is what happens when availability of communication drops? Redundancy of communication

13 See for instance DARPA's Guaranteeing AI Robustness Against Deception (GARD) program, https://www.darpa.mil/program/guaranteeing-ai-robustness-against-deception. At TNO research is being conducted on adversarial AI as well (counter AI and counter-counter AI).

lines to critical sensors and processes seems necessary.

Of course, technically facilitating meaningful human control per task is necessary too. In essence Athena is not just a technical system, but a socio technical system. We therefore propose that Athena comes with variable autonomy—dynamically adjustable levels of autonomy—as a means of ensuring meaningful human control by satisfying all three core values commonly advocated in ethical guidelines: accountability, responsibility, and transparency.[14] And human operators should be able to intervene at probably each stage. Athena is not HAL2000.

One of the most intriguing architectural questions is, whether Athena is a system of systems in your network, or a distributed 'environment', perhaps (partly) commercial Cloud based, perhaps partly hosted by a trusted partner. And could it be entirely outsourced, as a managed service? This would stress the importance of connectivity even more. And the importance of trust too, given the high level of access to networks and data![15] This is even more so when Athena autonomously shares data with other organisations.

Ethical / legal / societal considerations

If an Athena-like system can technically be designed and built, there are other, non-technical considerations that will impact the shape and size of the system. Like any business-critical AI application, it will be important to be able to understand how the system works. Explainable AI and accountability are necessary for human operators to understand how the system comes to its decision. Especially when system decisions (or advice for human operators) impact the physical environment or conflict with other interests of third parties.

Ethical

From an ethical point of view, explicitly embedding values in the system and provide accountability for the application of those values, is necessary. This is important for all sorts of autonomous systems, but who designs those values and how?! It should be a broad process involving experts, policy makers, civil society et cetera.[16] The more societal impact, the broader the consultation process. However, in digital technology the default was for a long time: the developers building the system incorporated their values implicitly (and probably unknowingly).

To Athena, the business and societal context of its specific implementation is important. It does make a difference whether it is deployed in, say, a hospital or a government agency.

Legal

Autonomous cyber security engines or systems require a lot of data to operate. Part of that data will be about people, like users and clients. Thus, privacy concerns might arise. Where to store[17] that data and how long to store?

The same goes for sensitive data, for instance about trade secrets or in military networks. Where to store and how to share and process classified data, even within the organisation, can be troublesome, and can render the autonomous concept unworkable or at least very inefficient, requiring a lot of human oversight. Although this is a 'man-made' legal issue, like in privacy the rules here protect real values and real lives. Perhaps surprisingly, autonomous security might be the hardest to implement in just those situations where the cyber risks are high enough to justify the investments...

A specific legal concern arises when the autonomous hack-back (if chosen to implement it as such) leads to direct or indirect impact outside the own network (and thus organisation). A simple example is taking down a command and control (C2) server of a botnet or the whole botnet to stop a distributed denial-of service (DDoS) operation against one's systems.

14 Methnani L, Aler Tubella A, Dignum V and Theodorou A (2021) Let Me Take Over: Variable Autonomy for Meaningful Human Control. Front. Artif. Intell. 4:737072. doi: 10.3389/frai.2021.737072

15 After citing Sun Tzu, we might as well involve Macchiavelli, who was critical about hiring armies. In his view you can only trust your own troops to be loyal and prepared to fight.

16 See for instance Ilse Verdiesen and Virginia Dignum, Value elicitation on a scenario of autonomous weapon system deployment: a qualitative study based on the value deliberation process, Springer, 2022.

17 EU regulation -> store in EU. Source?!

Societal considerations

The first societal consideration concerns one of the problems autonomous security should solve: the shortage of skilled personnel. Autonomous security is a complex mix of technologies, that draws on cyber security and another scarce discipline, namely AI. It is not clear how we as a society can fill those gaps in the short term. Finding synergy by multinational cooperation can be a start.

A second concern: it seems fair to assume that primarily large high-risk organisations are able and willing to make the necessarily investments, as the implementation of autonomous security would be quite costly. This widens the digital security gap only further, in which small organisations like for instance SMEs, healthcare and local government become more vulnerable by comparison.

An alternative approach could be providing autonomous security as an external managed service (cloud based or otherwise). This does bring its own considerations. We already mentioned the importance of trust and connectivity. And who will design, develop, train, configure, maintain, and sell these autonomous systems? A scenario looms in which American big tech cloud providers pocket this market too, since it will be a very complex machine to build, requiring a lot of expertise and substantial (risk seeking) funding and basically, mass. This not-unlikely development will conflict with the expressed need in Europe for more strategic autonomy in the digital domain. The European knowledge base seems up to the task, so we should aim to leverage it into healthy economic activity.

Conclusions

From a cyber resilience point of view, autonomous security already provides benefits in terms of speed, scale, and accuracy. Today, it offers added value to the existing security chain, from prevention to response and recovery, bolstering cyber resilience considerably, and helping to resolve existing human capital issues. Tomorrow however, a system like Athena might become the only way to fend off AI-attackers.

All-in-all, the proposed Athena-machine will be a very complex autonomous system of systems, requiring amongst others a lot of input, processing power and a clear concept and decision-making framework for response. Continuous learning should be an integral part of autonomy, preferably with other Athena-implementations. However, autonomy without automated (re-)configuration can still be highly useful. And as we've pointed out, there will be trade-offs too to meet technical, ethical, legal, and societal considerations.

The codename Athena is not by accident godlike. Autonomous cyber security requires omnipresence and omniscience (though not omnipotence as we discussed). Implementation to its fullest extent is not only very complex and vulnerable, but also raises numerous ethical, legal, and societal dilemma's. However, elements of Athena are valuable to substantially increase cyber resilience of organisations and society, supplemented with other approaches like "affordable security" (in multiple ways) and increased ease of use of security technology. More incremental is the continuing efforts on automated security in every chain, expanding towards technical architectures.. Lastly, in essence cyber security exists by grace of vulnerable IT systems, including people, processes and networks. If everything in this system worked flawlessly, there would be a much smaller need for cyber security innovations.[18] So why not address that more directly? Reducing inherent vulnerability substantially, will probably require a complete redesign of IT architectures as sociotechnical systems and is thus not very likely to happen in the short run or on a large scale. Part of the problem is the ever-increasing complexity of IT, to which cyber security technology might only add. However, reducing vulnerability and complexity of IT is obviously beneficial in many ways. So, starting with the cause in mind, this is also an interesting line of thought to explore.

18 Of course, even with perfect software there could still be cybercrime, with attack vectors like phishing, stealing user credentials. This would still require a certain level of protection.

Recommendations

Various technical, organisational, and ethical/legal/societal aspects need to be addressed in order to achieve anything like Athena. If it were to contribute to a more secure society, what should be the first steps to take towards autonomous security?

A first no regret step would be improving the technology for automated dynamic IT / OT / IoT asset discovery. This is also relevant for other topics like efficient IT management, crypto agility, and automated security.

Another key notion is that reasoning over cyber security requires modelling the cyber security space first. Elements to be modelled include IT infrastructures, cyber-attack vectors and responsive courses of action. Efforts to build such models lay a foundation for Athena. The Athena concept explained here is high level and deserves more detail. Design options and considerations should be explored by combinations of universities, RTO's, industry, and end users. More elaborate concepts could trigger more concrete Ethical, Legal and Societal Aspects (ELSA) discussions and lead to meaningful and well-scoped proofs of concept and experiments. This should merge with on-going automated security research & development, as this is aimed at part of the solution.

For ethical, legal, and societal aspects, we should also follow closely the discussions taking place in the AI arena, but also that of autonomous weapons (given the potential impact of cyber actions on the physical domain).

More specifically, lessons on effective human-machine interaction and meaningful control from other domains to the cyber domain should be considered.

Like for any AI application, data is everything. We need to start assembling datasets to provide a learning environment for autonomous security.

Given the implications for our national and European strategic autonomy, public intervention is required. Governments can initiate and stimulate R&D on autonomous security in various ways, utilizing existing structures and instruments already available to them. In this, given the sheer size and complexity of (semi) autonomous security, European cooperation is not just desirable, but a necessity.

If cyber security operators indeed find themselves outsmarted by AI soon, they will not be able to walk away from the chess board, like Kasparov in 1997. Preparing for the end of the age of AI-competitive humans means accepting that this time, losing is not an option.

Karin Bosch, Frank Fransen,
Patrick de Graaf, Dimitri Hehanussa,
Rick van der Kleij, Bert Jan te Paske,
Berry Vetjens and Reinder Wolthuis

Contact
Berry Vetjens

Director market unit ICT,
Strategy & Policy (ISP)

berry.vetjens@tno.nl
+31 6 53 44 14 54
https://www.linkedin.com/in/vetjens

TNO innovation for life

tno.nl/autocyber