2022 Richard S. Schultz '60 Fellow Research Report

Elaina Latino

"An Analysis of the Current Uses of AI and Why There is a Need for the Subfield of AI Forensics in Warfare"

Contents

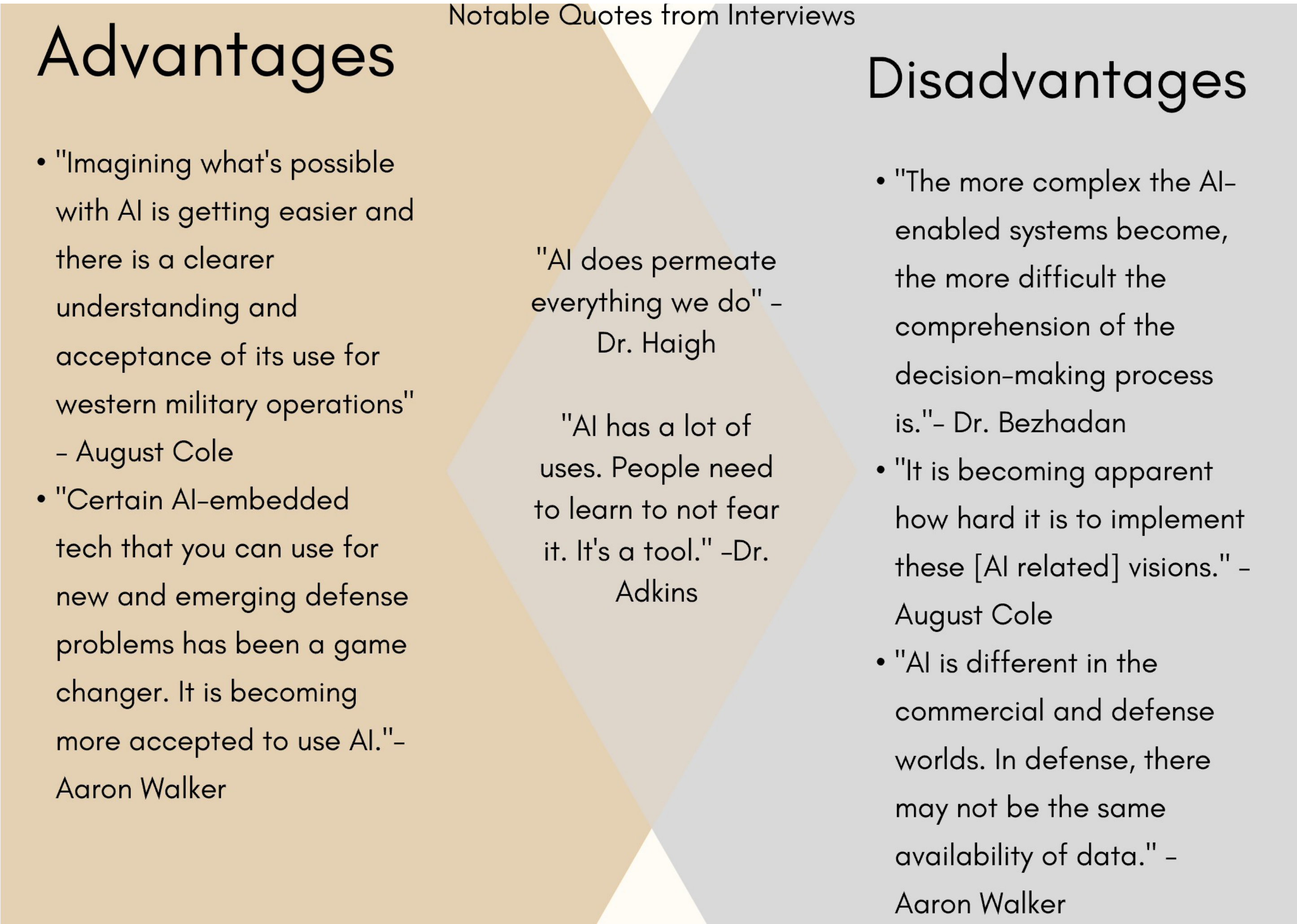
Abstract

Advances in artificial intelligence are leading to a new era in modern warfare. However, this new era will also bring AI-related accidents, criminal activity, and the need for expert evaluation of evidence to determine what went wrong, and what can be done to deter it in the future. Specialists with expertise in both artificial intelligence and digital forensics will be needed to tackle this new aspect of warfare. The complexity of artificial intelligence is one factor that makes it hard to investigate, and there will be a need for different experts who can evaluate unique subsets of AI Forensics.

Keywords: Artificial Intelligence, Digital Forensics, AI Forensics, AI warfare, UAVs, Explainable AI

Introduction

Although we are not living in Westworld where robots are indistinguishable from humans, we are in a time where we are surrounded by artificial intelligence (AI). From the curated ads on your phone to semi-autonomous cars driving on the road, AI is all around us. These developments cannot be contained (Scharre, 2018). The widespread usage of AI affects both civilian and military entities. In fact, some say artificial intelligence is the future of warfare (Tangredi & Galdorisi, 2021). However, the AI revolution brings with it new risks (Jensen, et al., 2019).



Notable Quotes from Interviews

Advantages

- "Imagining what's possible with AI is getting easier and there is a clearer understanding and acceptance of its use for western military operations" - August Cole
- "Certain AI-embedded tech that you can use for new and emerging defense problems has been a game changer. It is becoming more accepted to use AI."- Aaron Walker

"AI does permeate everything we do" - Dr. Haigh

"AI has a lot of uses. People need to learn to not fear it. It's a tool." -Dr. Adkins

Disadvantages

- "The more complex the AI-enabled systems become, the more difficult the comprehension of the decision-making process is."- Dr. Bezhadan
- "It is becoming apparent how hard it is to implement these [AI related] visions." - August Cole
- "AI is different in the commercial and defense worlds. In defense, there may not be the same availability of data." - Aaron Walker

Figure 1. Advantages and Disadvantages of Artificial Intelligence (AI). Experts in AI, forensics, and both fields, point out the pros and cons of incorporating AI into more systems (Latino, 2022).

The birth of artificial intelligence can be traced back to a workshop at Dartmouth College in the summer of 1955, organized by John McCarthy who later founded the Stanford AI Lab (McCarthy, J., et al, 1955). The lofty goal of the workshop was to determine if "every aspect of learning or any other feature of intelligence can be so precisely described that a machine can be made to simulate it" (Liang, 2020). Some exciting results from that workshop were programs that played chess and others that proved theorems (McCarthy, J., et al, 1955). There was overwhelming optimism about this new aspect of computer science:

"Machines will be capable, within twenty years of doing any work a man can do."
–Herbert Simon (Simon, 1960)

"Within ten years the problems of artificial intelligence will be substantially solved."
–Marvin Minsky (Minsky, 1967)

"I visualize a time when we will be to robots what dogs are to humans, and I'm rooting for the machines."
–Claude Shannon
(Omni Magazine, 1987)

Unfortunately, there were underwhelming results. Artificial intelligence research entered a quiet period (Allen, 2001). The implications of the early era contained two problems: limited computational power and limited access to information (Liang, 2020).

AI research that continued throughout the 1970s and 1980s had more of a narrow focus (Allen, 2001). Some applications impacted the industry significantly. For example, in 1989 applied convolutional neural networks were able to recognize handwritten digits, a capability ultimately used by the US Postal Service (Liang, 2020). The topic of deep learning made strides in the

1990s and continues until today (IBM Cloud Education, 2020). AI is now prevalent in all places from the phone in your hand to the newest high tech military weaponry (Haigh, 2022).

Staying technologically advanced is a critical way to stay ahead of adversaries, and AI is helping to do that. As AI is increasingly used in military systems, it is inevitable there will be instances where these systems fail. Whether accidental or criminal, these failures will warrant a technical investigation that produces legally acceptable and scientifically grounded findings on the causes – in other words, a forensic examination. Unfortunately, despite the proliferation of AI in civilian and military contexts and the subsequent incidents that are likely, AI forensics is not receiving the attention needed. Even with the expectation and occurrences of defects in AI-embedded systems, the idea of AI Forensics was only recently introduced (Baggili & Behzadan, 2019). Having the proper forensic tools available to examine systems used or involved in a military or civilian incident is also vital. New AI-embedded tools are being used to help process digital forensics investigations (Li, 2019). However, there is not much research investigating the forensics of AI itself, many of the tools needed for it are yet to be discovered and, overall, AI forensics is an area that is not well explored (Baggili & Behzadan, 2019). The next sections provide background information about digital forensics and artificial intelligence to better clarify these issues.

Background

What is Digital Forensics?

Digital forensics, also referred to as cyber forensics, was developed over time to deal with cyber-attacks and other incidents involving computer technology (Luciano, et al, 2018). As such incidents and attacks occurred there became a specialized need to produce legal and scientific findings based on evidence found in computer hardware, software, networks, and data storage (Yaacoub, et al., 2022). Simply put, digital forensics is the intersection of criminology, computer science, and law. It takes the "best tools out of these three domains and uses them in an interdisciplinary fashion to solve crime" (Adkins, 2022).

Forensic methods are applied to sources of potential evidence from an incident to learn and factually support a description of what happened (Pande, 2016). Because those methods and their results might be used in a court of law, they must also meet standards according (in the United States) to the Federal Rules of Evidence (Yaacoub, et al., 2022). Among those standards are a set known as the Daubert Standards (Fig. 2); their purpose is to assure the validity of scientific evidence used as the basis for expert testimony (Daubert v. Merrell Dow Pharmaceuticals, Inc., 1993; "Daubert Standard", n.d.). Digital forensic methods are no exception; they, too, must meet the Daubert tests (Brunty, 2011). Any investigation into an incident involves forensics and any such results may end up in a court of law, whether or not that is anticipated at the start of the investigation (Yaacoub, et al., 2022).

To be used in support of an expert opinion, any scientific results used as evidence in any investigation must meet the same standards (as shown in figure 2), but of course may use different methods for the analysis. The discipline of digital forensics has many sub-specialties and a vast array of methods due to the variety of digital-evidence sources and their critical differences–from mobile phones to emails to networks (Yaacoub, et al., 2022). All forensics follows a similar general process: identify potential evidence, collect it, preserve it, analyze it, and report on it (all while maintaining chain of custody) (e.g. DFRWS, 2001; Carrier and Spafford, 2004; Easttom, 2022). For example, readers with a military background may already be familiar with a digital-forensics technique used in Afghanistan to "exploit intelligence"– Document and Media Exploitation, or DOMEX (Sammons, 2015). With DOMEX, by forensically collecting and examining devices and digital media seized on the battlefield, examiners can uncover valuable information about adversaries. This information can then be used to generate actionable intelligence to improve battlefield efficiency and effectiveness (Sammons, 2015). As the use of computers, computerized devices, and embedded AI on the battlefield continues to increase, so too will the need for appropriate digital forensics methods and expertise to find, extract, and disseminate this information.

Different subfields of forensics exist because different evidence requires distinctly different scientific bases, tools, and forms of analysis. Digital forensics is no different and itself

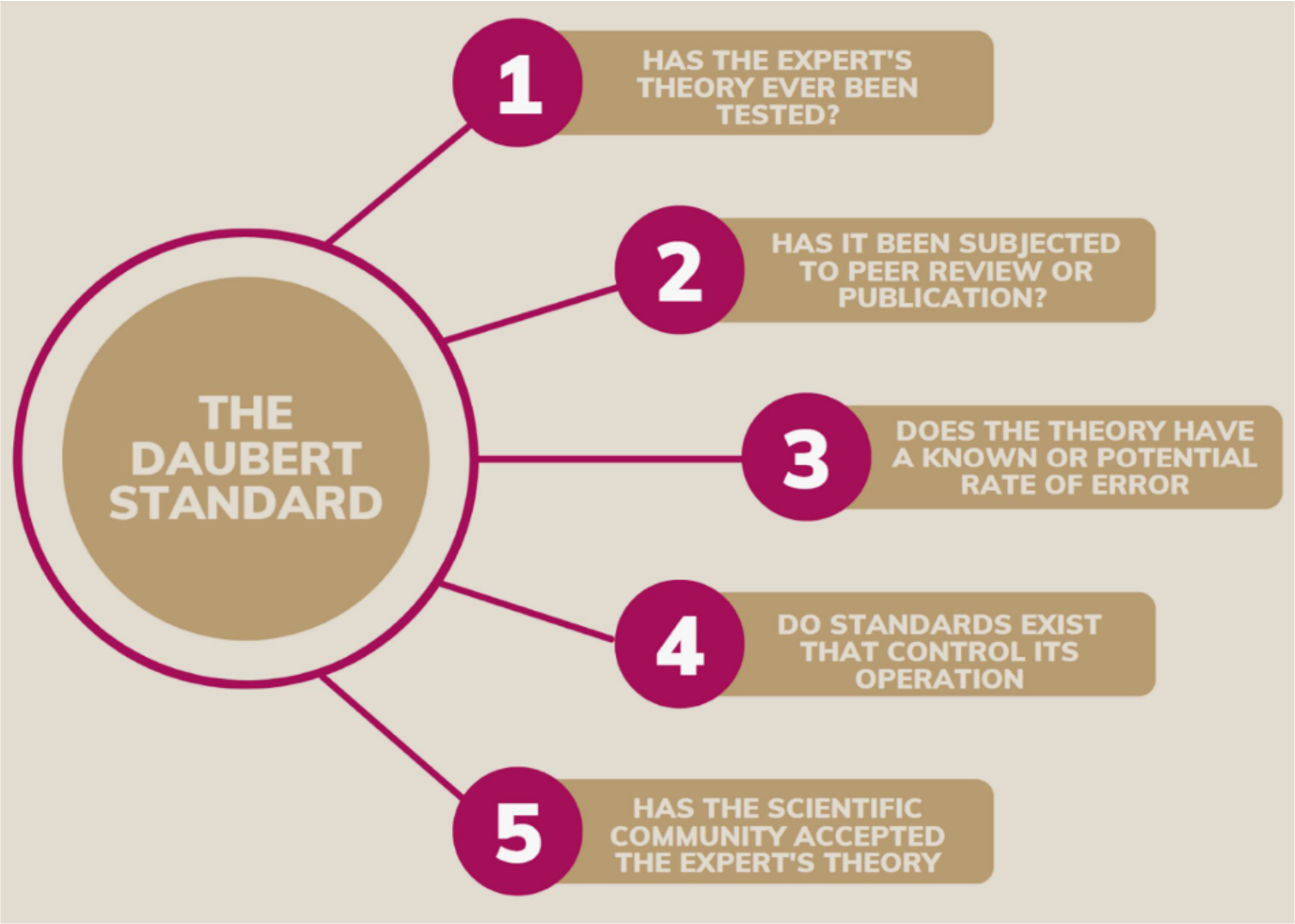

THE DAUBERT STANDARD
1
HAS THE EXPERT'S THEORY EVER BEEN TESTED?
2
HAS IT BEEN SUBJECTED TO PEER REVIEW OR PUBLICATION?
3
DOES THE THEORY HAVE A KNOWN OR POTENTIAL RATE OF ERROR
4
DO STANDARDS EXIST THAT CONTROL ITS OPERATION
5
HAS THE SCIENTIFIC COMMUNITY ACCEPTED THE EXPERT'S THEORY

Figure 2. The Daubert Standard Requirements for Scientific Evidence. According to the Federal Rules of Evidence (e.g see Yaacoub, et al., 2022), for scientific evidence to be used in support of expert testimony it must meet certain tests. It needs to be more than generally accepted in the scientific community – it needs to have been tested, with known or potential error rates and standards, and subjected to peer review or publication. These are known as the Daubert Standard or the Daubert tests ("Daubert Standard", 2022, redrawn by Latino, 2022)

has different complex subfields with different components, requiring distinct tools and forms of analysis within each such subfield (Yaacoub, et al., 2022). For example, mobile devices and computer networks, although both digital systems that interact, are different in many ways. Mobile devices can use a cellular network where computers need WiFi (Bouchrika, 2022). Even so, information from a mobile device, flowing over a cellular network, will at some point likely flow through a computer network that may (or may not) utilize WiFi (Bouchrika, 2022). Depending on the needs of an investigation, a forensic examiner might need to be experienced primarily with cellphones and smartphones, with information retrieved via subpoena from a cellular carrier, or with the methods needed for the various places where potential evidence might exist elsewhere in the larger, related computer network (as shown in Fig. 3) (Yaacoub, et al., 2022).

Consider another example, something as simple as computerized data-storage media. Earlier-generation computers used mechanical hard drives; modern ones and mobile devices now typically use solid-state memory (Sammons, 2015; Easttom, 2022). Each media type presents nuanced differences, challenges, and collection protocols that a forensic investigator must understand and correctly apply to reliably extract and preserve evidential data (Brunty, 2011). They must not only collect and preserve the evidence appropriately; they must be knowledgeable about the limits of the collection and further analyses (Yaacoub, et al., 2022).

Investigational needs can vary according to incident type and situational demand as well as device or technology. Where geographic location is important, mobile-device GPS and tracking-systems data would be vital; network evidence

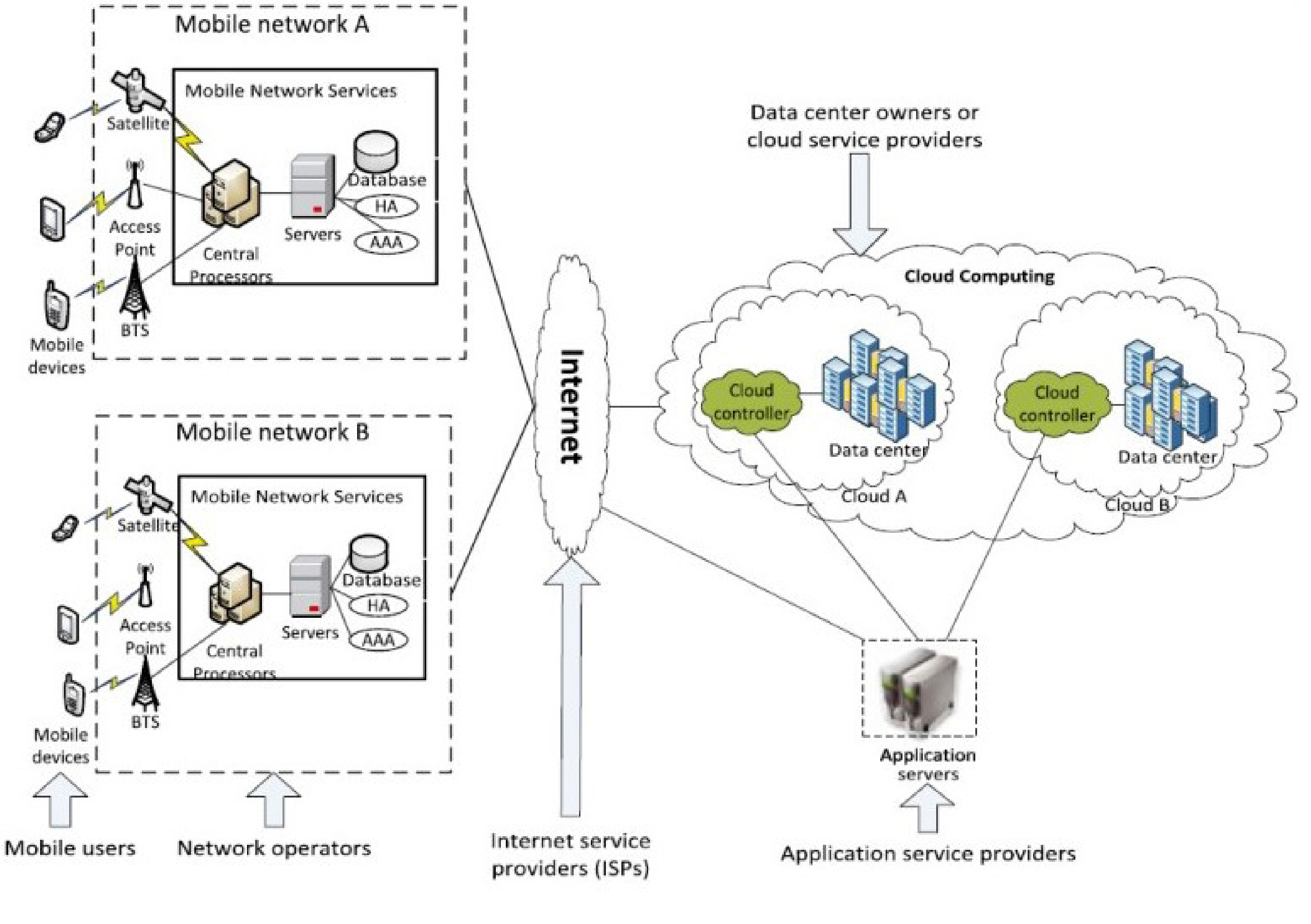


Figure 3. Mobile Cloud Computing Architecture of Network. The graphic highlights how complex the path for information is, and the many devices that might be involved in running to ground evidence for an investigation. (Gloukhovtsev, 2020)

may not be as important (Sammons, 2015; Easttom, 2022). Establishing a network of known associates may focus primarily on email and social-networking application evidence and less so on other types of evidence (Luciano, et al. 2018). Therefore, even though the basic forensic process and requirements for digital forensics, like the need for validated tools and quality assurance, may be the same, the forensic evaluation of each type of digital technology and its use can present unique challenges and require diverse tools and techniques (Sammons, 2015 & Pande, 2016).

Artificial intelligence is itself a type of digital computer system. In the same way each other subset of computer systems has common forensic requirements and yet demands specialized knowledge, skills, and abilities, AI introduces its own set of differences. AI-embedded technology will require not only specialized knowledge, skills, and abilities, but specialized tools for forensic investigations. The next section introduces AI and provides an overview of some of the key differences.

What is Artificial Intelligence?

IBM Cloud provides a simple yet compelling definition of AI: “Artificial intelligence leverages computers and machines to mimic the problem-solving and decision-making capabilities of the human mind” (IBM Cloud Education, 2020). There is a wide array of technology under the umbrella phrase “artificial intelligence.”

One way to think of it is as a Russian nesting doll (Adkins, 2022; see Fig. 4). Artificial intelligence – the generic parent concept – is the outermost “doll.” However, within that are more detailed sub-specialties, such as machine learning, deep learning, and more complex integrations of multiple methods. Machine learning algorithms can be trained to identify and collect valuable information, through a series of tests, with the help of human input. In the same way a toddler learns that falling down hurts and should be avoided, by falling a few times; a machine learning algorithm is trained to understand certain tasks by running through various scenarios, failing a few times, and being corrected with each trial (Haigh, 2022). Image recognition, such as medical image analysis, utilizes machine learning algorithms (IBM Cloud Education, 2020). Deep learning is said to imitate the way humans learn, without the need for human interaction (Adkins, 2022). Instead of relying on human input, deep learning utilizes neural networks, or systems modeled on the human brain, to learn (IBM Cloud Education, 2020). Personal recommendations seen on apps like Netflix or YouTube are produced by deep learning systems (Steck, et al., 2021). More complex applications include the use of AI in autonomous vehicles (Lee, 2019).

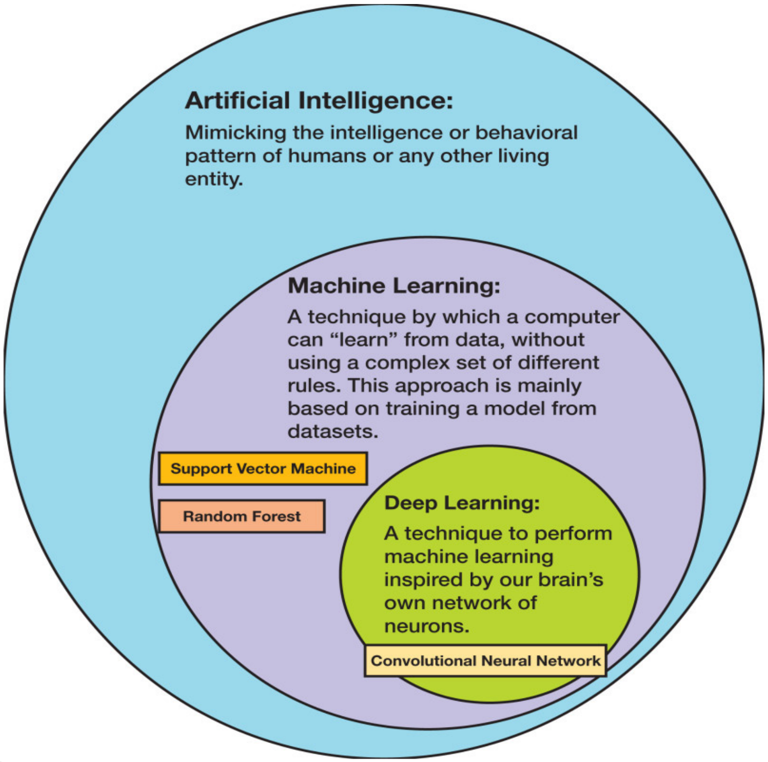

Artificial Intelligence:
Mimicking the intelligence or behavioral pattern of humans or any other living entity.
Machine Learning:
A technique by which a computer can “learn” from data, without using a complex set of different rules. This approach is mainly based on training a model from datasets.
Support Vector Machine
Random Forest
Deep Learning:
A technique to perform machine learning inspired by our brain’s own network of neurons.
Convolutional Neural Network

Figure 4. The Layers of AI. This graphic depicts the different layers of artificial intelligence. Artificial intelligence is a more general term that encapsulates many layers and parts. (Zhao and Krauze, 2021).

Common Uses of AI

Virtual Assistants

So-called "virtual assistants" such as Siri, Google Home, and Amazon's Alexa utilize AI. Although they are distinct systems, they all utilize comparable algorithms to complete a similar task. Machine learning algorithms in the form of Natural Language Processing (NLP) help 'virtual assistant' software gather the required data to be processed to retrieve the requested information (e.g. see Gebhart, 2020; IBM Cloud Education, 2020; Eckel, 2021; Steck, et al., 2021; Fingas, 2022; and, Wang, 2022). NLP is yet another subfield of machine learning that makes it possible for computers to comprehend and analyze human language (IBM Cloud Education, 2020). However, virtual assistants can do much more than simply answer a question such as "What is the weather?" Rather than simply identifying a need for information and retrieving it, they are expanding to include identifying patterns of actions to perform and accomplish a specific task, such as "make a bank transfer" (Gran, 2022). Virtual assistants can go beyond even this and help individuals maintain a schedule for their bank payments as well as simplify the steps needed to make the transfer (Gran, 2022). Virtual assistants are just one use of artificial intelligence. While they show how versatile AI can be, more complex applications of it are pervasively in use.

Autonomous Vehicles

The road to self-driving cars has advanced considerably over the past 15 years (Fig. 5). While autonomous systems such as Tesla's Autopilot feature have captured the attention of headlines (Lee, 2019), autonomous systems have been utilized in planes and trains for years as well. Dr. Karen Haigh, Consultant for Cognitive Electronic Warfare and Embedded AI/ML said:

"Boeing (aerospace company) brags about the fact that on any given flight, a human is only responsible for 7 minutes of flight, everything else is done by AI. Airbus laughs at Boeing. What do you mean 7 minutes, [our pilots are only responsible for] 3 and a half minutes" (Haigh, 2022).

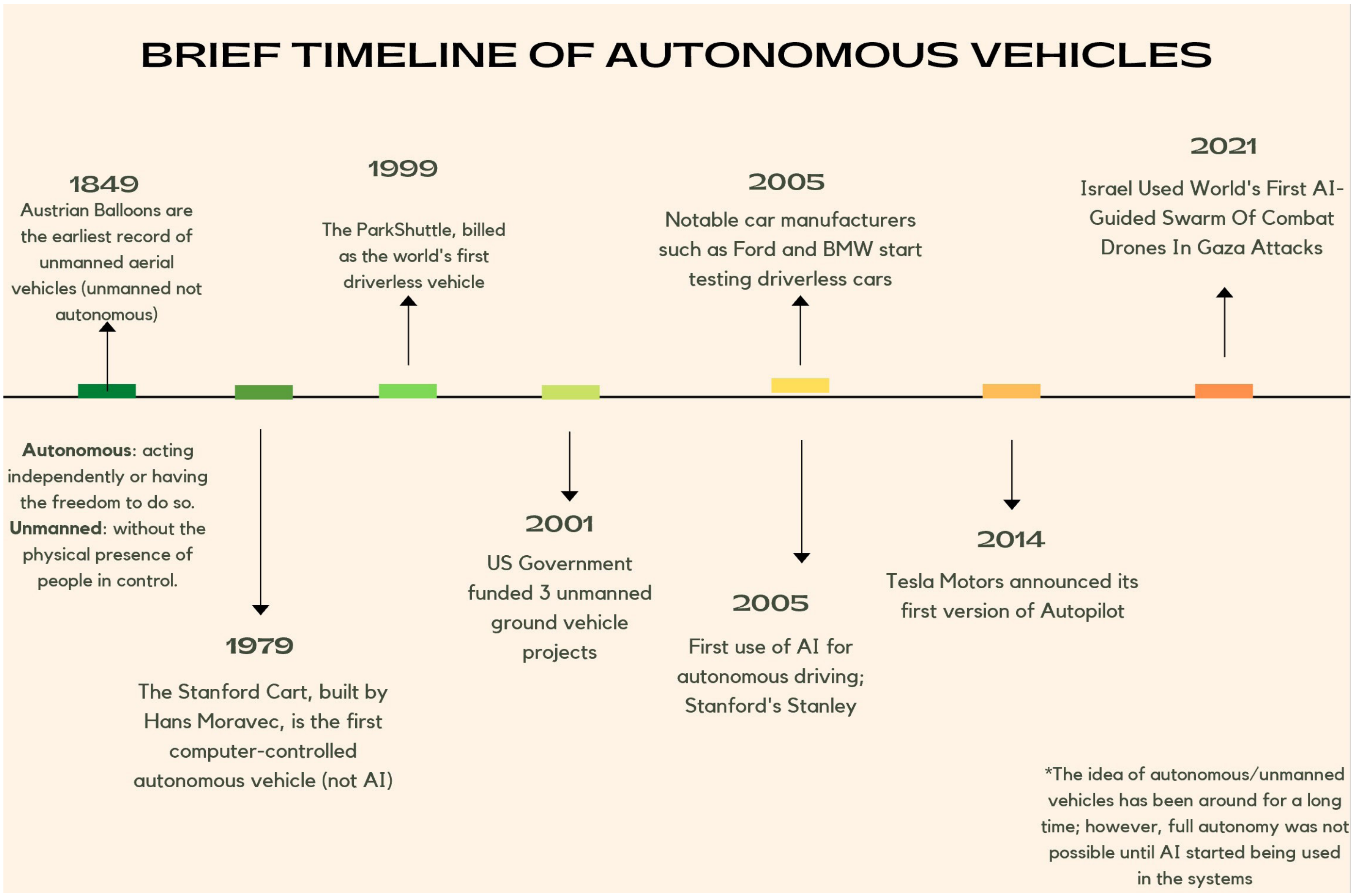

BRIEF TIMELINE OF AUTONOMOUS VEHICLES
1849
Austrian Balloons are the earliest record of unmanned aerial vehicles (unmanned not autonomous)
1999
The ParkShuttle, billed as the world's first driverless vehicle
2005
Notable car manufacturers such as Ford and BMW start testing driverless cars
2021
Israel Used World's First AI-Guided Swarm Of Combat Drones In Gaza Attacks
Autonomous: acting independently or having the freedom to do so.
Unmanned: without the physical presence of people in control.
1979
The Stanford Cart, built by Hans Moravec, is the first computer-controlled autonomous vehicle (not AI)
2001
US Government funded 3 unmanned ground vehicle projects
2005
First use of AI for autonomous driving; Stanford's Stanley
2014
Tesla Motors announced its first version of Autopilot
*The idea of autonomous/unmanned vehicles has been around for a long time; however, full autonomy was not possible until AI started being used in the systems

Figure 5. Timeline of Autonomous Vehicles. Achieving autonomy has been a longtime goal for inventors. With the help of artificial intelligence, true autonomy could be possible. (Daly, 2022, Editorial, 2021 & Malewar, 2022. Redrawn by Latino, 2022)

Virtual assistants and autonomous vehicles are only becoming more common. As a high-functioning and newly adaptive technology, AI could be seen as invasive. Due to the burgeoning presence of AI in systems and technological items, there is high importance for a field of individuals who understand how to deal with failures in AI systems.

Why these are important

Although these systems may be seen as invasive, their use is already too heavily relied on to be controlled (Haigh, 2022). Are virtual assistants and autonomous systems safe, or do they present security risks? Since the 'assistants' know and store valuable information (including details of behaviors), a breach caused by (or related to the use of) one could lead to the exposure of confidential data (Yampolskiy, 2016). When an autonomous vehicle malfunctions or causes an accident (Scherer, 2015), who or what was responsible? In terms of dealing with the use of AI systems, many complex legal questions have been asked and discussed, but few laws or regulations have been passed (Haney, 2020). Determining what happened by forensically evaluating the evidence of any technological system involved is crucial so an organization or a court of law can attribute responsibility. Doing that requires detailed knowledge of how such systems – including any embedded or assistive AI – function.

How AI is currently being used in the military

Artificial intelligence is also being used increasingly to benefit militaries around the world. From surveillance systems to unmanned aerial vehicles (UAVs), AI is significantly changing modern warfare. For example, satellite data analysis and filtering have advanced greatly due to machine learning algorithms that can handle substantial amounts of data (Haigh, 2022). Surveillance systems, which utilize a subset of AI called facial recognition, are used to gather information on people or militaries of interest (Haney, 2020). AI enables military vessels such as robotic ships and submarine prototypes to navigate without the need for a sailor (Tangredi & Galdorisi, 2021). Another military application of AI is its incorporation not into weapons systems, but the weapons themselves. Smart munitions and unmanned aerial vehicles are two examples.

Smart Munitions

While the use of AI in a weapon system can inflict greater or more severe damage, smart weapon systems can also be used to minimize unnecessary damage or unintended suffering. Dr. Daniel Ragsdale, Former Assistant Director for Cyber in the Office of the DoD, emphasized many reasons we

should want to see AI implemented into military systems. With the right training, AI could better determine "if there is a military necessity to apply lethal force. [Militaries] could have a munition for example, able to make a decision based upon the logic that is now a part of the weapon to discern whether or not you should engage" (Ragsdale, 2022). Currently, smart munitions are being used in Ukraine (Ragsdale, 2022). Provided by the United States, High Mobility Artillery Rocket Systems (HIMARS) have been able to lead more precise attacks through satellite navigation (Lopez, 2022). Due to the long strike range, Ukraine has been able to precisely strike Russian targets while simultaneously reducing danger to nearby civilians (Khurshudyan & Khudov, 2022).

Lethal Autonomous Weapons – Unmanned Aerial Vehicles

Lethal autonomous weapons are no longer just the science fiction of RoboCop or other films. According to *Army of None* author Paul Scharre, unmanned aerial vehicles (UAVs, or simply "drones") are being used in about half (over 90) of the militaries around the world (Scharre, 2018). Drones can be used solely for surveillance, or they can carry ordinance for use if surveillance reveals an actionable target. According to Sharre (2018), 16 of the militaries using drones use armed ones (Scharre, 2018). One example is the Israeli Harpy Drone. The autonomous systems in the Harpy drone were developed to search for radar targets (Scharre, 2018). Considered a 'loitering' munition, the Harpy can fly to an area and wait ("loiter"), until the set target is found (Walker, 2021). Once the appropriate radiation pattern is detected, the munition, which carries explosives, will dive into its target (Scharre, 2018).

Loitering munitions such as the Harpy have already been said to have helped win a war. During the Armenian-Azerbaijan border conflict, investments by Azerbaijan in loitering munitions gave them a significant advantage over Armenia (which was reported to only have one model) (Walker, 2021). This advancement in autonomous weapon systems helped lead them to victory (Walker, 2021).

Information Gathering

Information and intelligence-gathering has come a long way since the times of spymaster George Washington and the Culper spycraft used in the Revolutionary War (Mount Vernon Ladies' Association of the Union [MVLAU], 2022). The current focus is now on computer hacking, driven – or at least supplemented heavily – by artificial intelligence (Haney, 2020). It has been stated "the nation that best collects and makes sense of data through AI-applications will increase its military power and realize long-term competitive advantages" (Jensen, et al., 2019). Cyber-attacks that seek to gain confidential information are on the rise due to individual countries' hunts for tactical advantage (Ragsdale, 2022).

Why these are important

These technological advancements provide significant benefits to military organizations. Some of these systems are in widespread and common use; some are still in their preliminary stages. However, over the upcoming years progress will continue to be made. Although not every country is currently utilizing these new AI-embedded systems, more and more will follow. If the right procedures and tools are not developed, AI could grow in unexpected ways. With the expanded use of AI, issues have already occurred.

AI Issues

According to Kroenke and Boyle, (2019), incident causes can be categorized as human error, natural disaster, and malicious activity. As much as we may attempt to avoid them, the first two types – accidents – will happen. As with other scientific advancements, humans are also "very, very good at finding ways to use things in negative and unpleasant ways… and whatever regulations we want to put in place as a society is not going to cover the creativity of" those with malicious intent (Haigh, 2022). Unintended outcomes or abuse of AI could jeopardize both digital security and physical security (Johnson, 2020).

Adversarial Attacks and Deception

It is important to remember that there is no system that is 100% secure. Systems will fail (Yampolskiy, 2016). Just as there have been advancements in digital, physical, and political systems because of AI, there are many security risks that are of utmost importance. Both in civilian and military scopes, major threats arise with the expanded use of AI and some issues have already been detected (Johnson, 2020).

One notable example of deception highlighted by Ragsdale (2022) is the “mislabeled panda” (Hao, 2020; see Fig. 6). Through image recognition, or the method to identify and extract key figures from images and use them as input, machine learning models are trained to recognize specific entities in photos (Hao, 2020). By adding a bit of noise, a picture of a panda was confidently labeled by a trained AI image-recognition system as a gibbon (Hao, 2020). In spycraft, misinformation has a long history (MVLAU, 2022), and the mislabeled panda example can be thought of as analogous to ‘digital misinformation’ (Hao, 2020). The deliberate introduction of noise, or dataset errors, into an opponent’s data gathering can be seen as a form of adversarial misinformation attack – an attempt to ‘poison’ data gathered by autonomous systems (Hao, 2020). From finance to the military, data poisoning of an AI system could damage digital, physical, and political security and cause serious harm.

Complex Environments

As opposed to deliberate deception, misinformation caused by misinterpretation of complex environments is a common issue. In real-world domains, machine-learning algorithms face rare scenarios the algorithms may not be trained to handle (Behzadan, 2022). In fact, by their very definition, rare events and scenarios are unlikely to be part of a training set and may even be difficult for humans to anticipate a need to include. In reference to autonomous cars, Dr. Behzadan gave a good example in which the sensors in autonomous cars were easily confused by graffiti on stop signs or on the backs of trucks (Behzadan, 2022). Confused AI, more specifically autopilots used in autonomous vehicles, have already caused (or at least *allowed*) death to occur (Hawkins, 2019 & Lee, 2019). In 2019, a driver of a Tesla Model 3 turned on autopilot. Ten seconds later, the vehicle drove into a semi-truck that was crossing in front of it, killing the driver (Lee, 2019). A situation like this is complex: the driver should have been attentive, as Teslas are only semi-autonomous, yet the AI systems should have detected the truck. Neither the driver nor the AI detected the truck or initiated evasive action. Who or what was at fault in this situation?

For Prussian military theorist Carl von Clausewitz, “war is the realm of chance” … since all information and assumptions are open to doubt, and with chance at work everywhere, the commander continually finds that things are

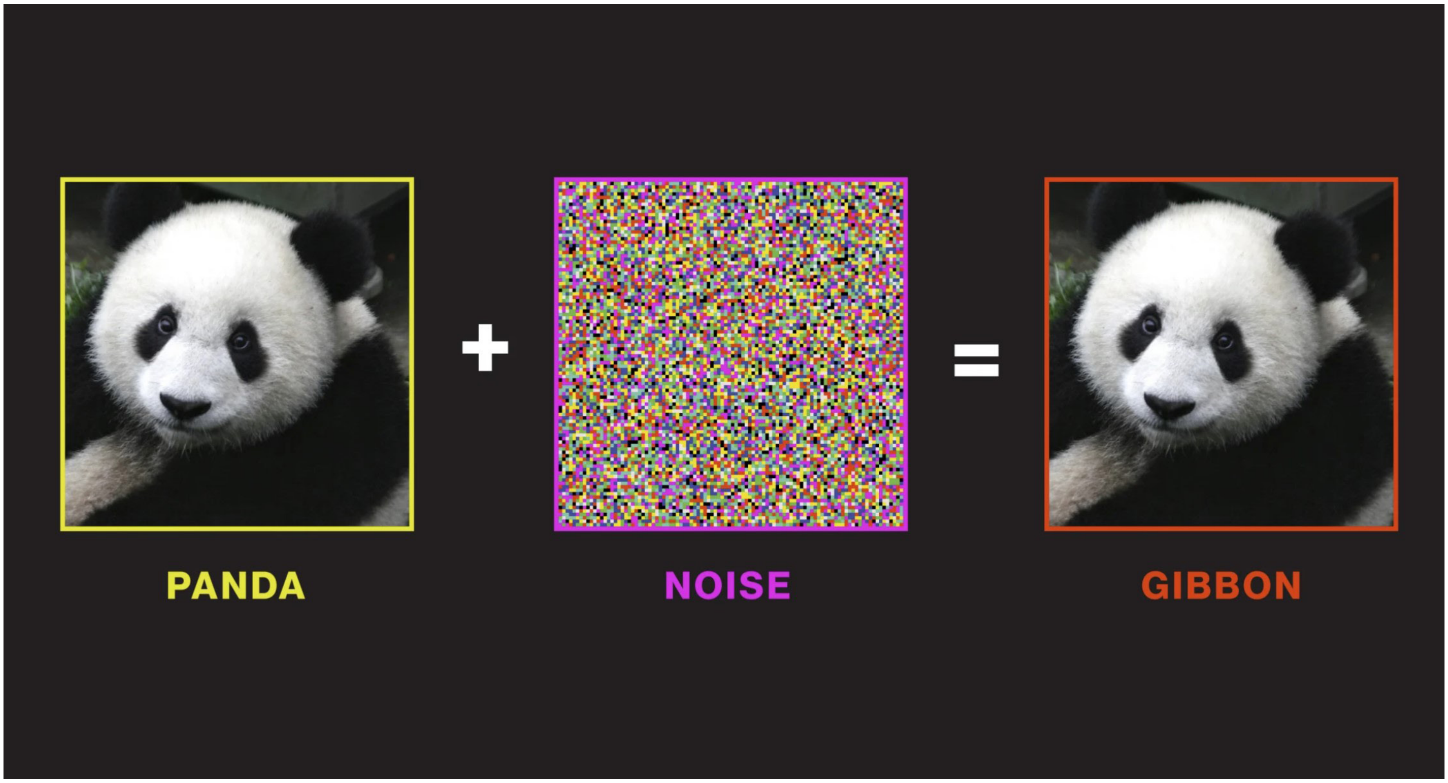


Figure 6. Data Poisoning Panda. Machine learning models are trained through repetitive tests. However, even well-trained models fail as the addition of noise can confuse the model. (Hao, 2020)

not as he[she] expected" (Tangredi & Galdorisi, 2021). Although there is no notable example of a failure like this happening yet in a war conflict, it is easy to imagine what could happen if an autonomous war vehicle or weapons system were brought into an environment filled with adversaries looking to deceive the device, or 'rare' circumstances and inputs. Confused AI could trigger events causing injury or loss of life, or poison tactical or strategic decision-making.

Explainable AI

Machine learning algorithms are decision-making software 'machines.' Yet AI processes are frequently referred to as either "black box" or "white box" (Jeong, 2020). When the decision-making process of a machine learning model cannot be explained, this is referred to as a "black box" process (Behzadan, 2022) – you cannot 'see inside it' to explain the results. When the "how" or "why" can be explained it is considered a "white box" process. To keep these systems and those who may be affected by their usage safe, it is necessary to be able to explain how it arrived at a given decision (Baggili & Behzadan, 2019). Explainable AI (or "XAI") aims to "enable human users to understand, appropriately trust, and effectively manage the emerging generation of artificially intelligent partners" (Turek, 2018). For forensic investigative purposes, all AI should also be explainable "white box" systems (Behzadan, 2022). However, they are not. "Given the amount of mystery, in terms of how different neural networks operate, it will be very hard to figure out at times, forensically why a system has done something" (August Cole, 2022).

Many people and organizations are trying to prioritize XAI. However, as noted by Cole (2022), explaining how AI systems arrive at an output or 'decision' is not simple. Furthermore, simplifying the systems so they are explainable may cause issues with performance results (Tangredi & Galdorisi, 2021). Consider the scenario where a simplified, explainable AI system, or restrained AI, combats an unrestrained AI system (Tangredi & Galdorisi, 2021). Which would win? Conventional theory suggests the unrestrained AI systems would be capable of using any means necessary, even (perhaps, especially) those beyond human anticipation, reasoning, or convention, and therefore have the edge (Clark & McLemore, 2019). By not having to take extra steps (or constraints) to be explainable, unrestrained systems are more efficient and effective in complex environments (Clark & McLemore, 2019). This presents an unenviable choice. Do people sacrifice explain-ability to improve the efficiency and effectiveness of the machines they use, or do they keep the machines explainable but with decreased capability?

How are these issues relevant to AI Forensics?

The purpose of a forensic investigation is to identify and use the available evidence, use valid scientific methods and procedures, and attempt to determine what happened, how, and why. However, with an issue like black-box AI, it may be difficult to understand why something went wrong. Data poisoning, like adding noise, is a simple yet effective way to confuse machine-learning algorithms. Forensic explanations often depend on a timeline of events (Pande, 2016). However, given complex and unknowable black-box processing, explaining an AI-related time sequence could be difficult. Forensics also seeks to help determine attribution – who or what was the root cause of an event or outcome. However, in the situation involving the Tesla, would the original programmers be at fault for not testing every scenario? Or would it be the driver for relying too heavily on a semi-autonomous system? Although this may not always be important, for a case that might be brought to court, explain-ability is key.

Other issues like data poisoning and deception have some solutions on the distant horizon. AIA Guard, "an end-to-end cybersecurity solution specifically designed against Artificial Intelligence Attacks" is one example ("AIA Guard"). However, this is just one solution under development, and these systems are not yet prepared for complex environments such as the battlefield. More forensic tools, like AIA Guard, need to be created with an emphasis on complex environments.

AI Forensics in Warfare

In current digital forensics investigations, analyzing computer and system access logs and tracking network traffic sources are vital parts of the investigation (Sammons, 2015; Easttom, 2022). The same need exists for an AI forensic investigation. It is important, especially regarding the military, to have this subfield to help create improved defense systems for future attacks. Each application domain, in this case AI forensics in warfare, has its own range of failures and because of this they each need domain-specific solutions (Johnson, 2020).

Many of the experts interviewed

for this paper mentioned how there is a lack of progress with the tools and advancements used in digital forensics. "Many of the tools and technics required for a thorough, scientific and legally accepted investigation are not yet developed," (Behzadan, 2022). Note this is just for general forensic purposes. With the ongoing applications and advancements to weapons and weapons systems, there especially needs to be more investments made in the forensics of military AI use.

With a field as complex as artificial intelligence, forensics investigations will be complex as well. Embedded in both hardware and software, contrasting AI issues may not be solvable with the same tools. Digital and physical systems may overlap. For example, autonomous drones utilize software, however, the software is embedded in a physical computing system, itself embedded in a flying drone carrying ordnance (Tangredi & Galdorisi, 2021). Failures in drones and failures in surveillance systems may have common, or vastly different, root causes, event timelines, and outcomes.

AI Forensic Environments

Over the course of researching this paper, many different AI-embedded systems were mentioned. In each and every one, failures can occur. Failures in physical systems such as unmanned aerial vehicles or smart munitions could lead to physical damage. For example, if a UAV were using facial recognition to track targets, then identified and fired on the wrong individual, an investigator would need to uncover how this mistake happened. If a financial or operational data-gathering system caused loss or harm, why did this occur? Was it accidental in nature, malicious, or environmentally caused? Was it a problem with the programming or the training sets? Was it a hardware malfunction? Was data poisoning used to pollute the AI pattern recognition or decision-making, as shown earlier with the

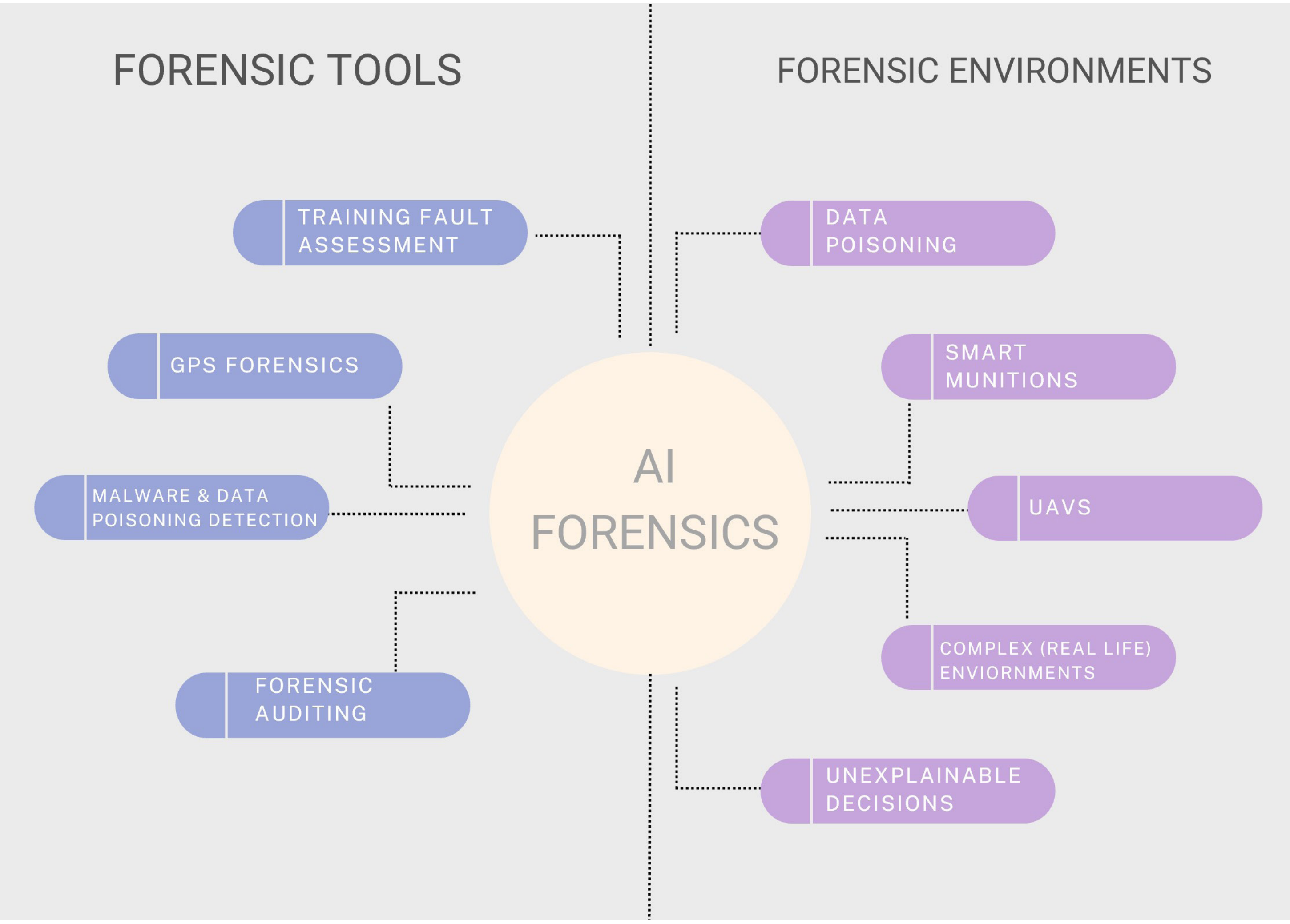


Figure 7. Tools and Environments involved in AI Forensics. The field of AI Forensics is a complex area field with different AI-embedded systems that will need to be investigated as well as different tools needed to help investigate. (Jeong, 2020, Yaacoub, 2022, & Ragsdale, 2022 redrawn by Latino, 2022)

panda picture? Or was there simply normal 'noise' that was such a 'rare' occurrence it had not been accounted for in the development of the AI and the system it was embedded in?

AI Forensic Tools

At the core of AI Forensics and the tools needed is the Daubert standard. As previously mentioned, this includes five criteria forensic findings must meet to be used in court. Among those criteria is a known or potential error rate (Brunty, 2011). Finding the "potential rate of error" will be extremely crucial for an AI forensics investigation. This relates directly to Explainable AI. To have a successful investigation one must be able to rely on the evidence, and the Daubert Standards are the current legal test of reliability.

As mentioned before, systems like AIA Guard are a step in the right direction when it comes to protection over malware and data poisoning in machine learning algorithms ("AIA Guard"). With technology like AI, itself designed to learn and improve, the only way to keep systems safe is with other AI-embedded systems. A common metaphor for the situation we find ourselves in is an 'arms race' or, more flippantly, a Catch-22 (Heller, 1961). Adversaries will attack. Being prepared for those attacks will be valuable. Since they are attacking with AI-enhanced systems, we need to defend with AI, creating more need for understanding how those systems were designed, how they function, and how they arrived at their outcomes. In short, more need for AI forensics.

For example, most AI systems will have (or have the capability to include) global positioning satellite (GPS) devices. The GPS system could potentially find the initial location of where and when something went wrong in the physical system. In reference to an incident involving something like an unmanned aerial vehicle, GPS could be extremely valuable. Analyzing the patterns could help investigators identify information in the environment at the time of a malfunction and begin to understand why the system reacted in a certain way. In the same way that criminal investigators may retrace the steps of the offender, AI forensics investigators should retrace the "steps" of the system.

Conclusion

Multiple sources, expert opinion, and common sense arrive at the same conclusion – "AI is permeating everything we do" (Haigh, 2022). Although the present and the near future of artificial intelligence and autonomous weapons systems may not look like how the movies portray them, they are becoming more common each day. AI has already played a leading role in deciding the outcome of one war and is routinely employed by nearly half the militaries in the world. The use of unmanned aerial vehicles will continue to become a standard, and we can expect AI to have an expanded role in the future of warfare. The current conflict in Ukraine shows just how vital these new technological advancements can be (Khurshudyan & Khudov, 2022).

History has shown us that incidents have and will continue to occur with all applications of technology – civilian or military. When it comes to AI-embedded systems, "AI will not correct human errors… it will simply learn them," ("AI Could Reduce Human Error Rate," 2018). When things go wrong it is human nature to want to know how, why, and (where appropriate) who is responsible. We do not wish to accept a "black box" explanation of 'incidents.' We may, in fact, be compelled by law, regulation, or international agreement not to do so. Out of operational and legal prudence, or necessity, we will need assured methods to find out what happened – in other words, we will need to be capable of forensically evaluating the evidence. Yet AI systems challenge our ability to do this. Furthermore, the development of AI forensics as a field, and AI forensic tools, woefully lags behind the proliferation of AI use in multiple domains.

Due to the rapid and widespread adoption of various AI approaches and their complexity, it is therefore important to develop the needed specialist forensic expertise and tools for commercial and military applications. If we do, just like with all new emerging branches of digital technology, a new field of forensics will emerge. AI is already in use all around us. Now is the time for the emergence of AI forensics.

Author Biography

Elaina Latino is from Atkinson, New Hampshire. She is currently a junior at Norwich University studying computer security and information assurance with a concentration in digital forensics. Although new to the field, her summer research presentation on AI Forensics has helped her engage with experts in both artificial intelligence and digital forensics. This opened new doors for her and sparked an interest in her future. Over the summer, Elaina had the chance to study abroad through Norwich's Maymester. The immersive class on cyber surveillance allowed her to explore new areas of computer safety in Germany. Elaina has a passion for digital forensics, but on the side, she also enjoys surfing while she is home for the summer as well as playing club field hockey while at school.

References

Adkins, Jonathan. "Digital Forensics and Deep Learning Interview." 2022.

"AIA Guard – Make Your AI Systems Safe & Reliable." *AIA Guard - Make Your AI Systems Safe & Reliable*, **https://aiaguard.com/#why**.

"AI Could Reduce Human Error Rate." Hello Future, Orange, 15 Oct. 2018, https://hellofuture.orange.com/en/ai-reduce-human-error-rate/.

Allen, Frederick. "The Myth of Artificial Intelligence." *AMERICAN HERITAGE*, American Heritage Publishing Co., 1 Sept. 2022, **https://www.americanheritage.com/myth-artificial-intelligence**.

Baggili, Ibrahim, and Vahid Behzadan. *Founding the Domain of AI Forensics*, 12 Nov. 2019, **https://doi.org/ Electrical** & Computer Engineering and Computer Science Faculty Publications. (B)

Behzadan, Vahid. "AI Forensics Interview ." 2022.

Bouchrika, Imed. "Mobile vs Desktop Usage Statistics for 2021/2022." *Research.com*, Research.com, 15 Sept. 2022, **https://research.com/software/mobile-vs-desktop-usage#:~:text=Another%20major%20difference%20between%20a%20mobile%20and%20a,a%20WiFi%20connection%20to%20allow%20for%20Internet%20access**.

Brunty, Josh. *Validation of Forensic Tools and Software: A Quick Guide for the Digital Forensic Examiner* , 3 Feb. 2011, **https://www.researchgate.net/profile/Josh-Brunty/publication/320808735_Validation_of_Forensic_Tools_and_Software_A_Quick_Guide_for_the_Digital_Forensic_Examiner/links/5e2f0643a6fdcc3096941501/Validation-of-Forensic-Tools-and-Software-A-Quick-Guide-for-the-Digital-Forensic-Examiner.pdf**.

Easttom, C. (2022). Digital Forensics, Investigation, and Response, 4th Edition. Jones & Bartlett Learning, Burlington, MA.

Eckel, E. (2021). "Apple's Siri: A cheat sheet". TechRepublic. **https://www.techrepublic.com/article/apples-siri-the-smart-persons-guide/**. Last visited 27 September 2022.

Carrier, Brian and Eugene Spafford (August, 2004). “An Event-Based Digital Forensic Investigation Framework”, Proc. Digital Forensic Research Conference (DFRWS) 2004, Baltimore, MD. **https://dfrws.org/wp-content/uploads/2019/06/2004_USA_paper-an_event-based_digital_forensic_investigation_framework.pdf**

Clark, Charles, and Connor McLemore. “The Devil You Know: Trust in Military Applications of Artificial Intelligence.” War on the Rocks, METAMORPHIC MEDIA., 23 Sept. 2019, **https://warontherocks.com/2019/09/the-devil-you-know-trust-in-military-applications-of-artificial-intelligence/**.

Cognitive Technologies Group. “Ai-Based Shunting Locomotive DAS Receives Declaration of Conformity.” *Global Railway Review*, 25 Jan. 2022, **https://www.globalrailwayreview.com/news/131433/ai-based-shunting-locomotive-declaration-conformity/.**

Cole, August. “Artificial Intelligence, Futurist Interview.” 2022.

Daly, David. “A Not-so-Short History of Unmanned Aerial Vehicles (UAV).” *Consortiq*, 7 Sept. 2022, **https://consortiq.com/uas-resources/short-history-unmanned-aerial-vehicles-uavs#:~:text=1849%20%E2%80%93%20The%20First%20Military%20Use%20of%20UAVs,to%20attack%20Venice%2C%20but%20they%20were%20mostly%20ineffective**.

“Daubert Standard.” Cornell Law School, Legal Information Institute, Sept. 2022, https://www.law.cornell.edu/wex/daubert_standard.

Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579 (1993)

Digital Forensics Research Workshop (August, 2001). “A Road Map for Digital Forensic Research”, *DFRWS Technical Report (DTR) T001-01*, Utica, NY. **http://old.dfrws.org/2001/dfrws-rm-final.pdf**

Editorial. “History of Autonomous Vehicles - Timeline.” *RoboticsBiz*, 1 July 2021, **https://roboticsbiz.com/history-of-autonomous-vehicles-timeline/#:~:text=History%20of%20autonomous%20vehicles%20%E2%80%93%20Timeline.%201938%20%E2%80%93,Remote%20controlled%20car%20tested%20at%20Ohio%20State%20University**.

“Forensic Definition & Meaning.” *Dictionary.com*, Dictionary.com, **https://www.dictionary.com/browse/forensic**. (T2)

Fingas, R. (2022). “How does Alexa work? The tech behind Amazon’s virtual assistant, explained.” Android Authority. **https://www.androidauthority.com/how-does-alexa-work-3209316/**. Last visited 27 September 2022.

Gebhart, A. (September, 2020). “Everything you want to know about Google Home.” CNET. **https://www.cnet.com/home/smart-home/everything-you-want-to-know-about-google-home/**. Last visited 27 September 2022.

Gloukhovtsev, Mikhail. “How 5G Transforms Cloud Computing .” *Dell Technologies* , Dell Inc, 2020, **https://education.dellemc.com/content/dam/dell-emc/documents/en-us/2020KS_Gloukhovtsev_How_5G_Transforms_Cloud_Computing.pdf**.

Gran, Ben. “How Banking Virtual Assistants Can Improve Your Banking Experience.” *Forbes*, Forbes Magazine, 5 Apr. 2022, **https://www.forbes.com/advisor/banking/banking-virtual-assistants/#:~:text=How%20Banking%20Virtual%20Assistants%20Can%20Improve%20Your%20Banking,Are%20Part%20of%20the%20Future%20of%20Banking%20**.

Haney, Brian Seamus. “Applied Artificial Intelligence in Modern Warfare & National Security Policy.” *SSRN Electronic Journal*, Winter. 2020, **https://doi.org/10.2139/ssrn.3454204**.

Haigh, Karen. “Artificial Intelligence and Cognitive Warfare Interview.” 2022.

Hao, Karen. “How We Might Protect Ourselves from Malicious Ai.” *MIT Technology Review*, MIT Technology Review, 2 Apr. 2020, **https://www.technologyreview.com/2019/05/19/135299/how-we-might-protect-ourselves-from-malicious-ai/**.

Hawkins, Andrew J. “Tesla’s Autopilot Was Engaged When Model 3 Crashed into Truck, Report States.” *The Verge*, The Verge, 16 May 2019, **https://www.theverge.com/2019/5/16/18627766/tesla-autopilot-fatal-crash-delray-florida-ntsb-model-3**.

Heller, Joseph. *Catch-22*. Simon & Schuster, 1961.

IBM Cloud Education. “What Is Artificial Intelligence (AI)?” *IBM Cloud Learn Hub*, 3 June 2020, **https://www.ibm.com/cloud/learn/what-is-artificial-intelligence**.

Jensen, Benjamin M, et al. “Algorithms at War: The Promise, Peril, and Limits of Artificial Intelligence.” *International Studies Review*, vol. 22, no. 3, 24 June 2019, pp. 526–550., **https://doi.org/10.1093/isr/viz025**.

Jeong, Doowon. “Artificial Intelligence Security Threat, Crime, and Forensics: Taxonomy and Open Issues.” *IEEE Access*, vol. 8, 7 Oct. 2020, pp. 184560–184574., **https://doi.org/10.1109/access.2020.3029280**.

Johnson, Bonnie. “Safety in AI-Enabled Warfare Decision Aids .” *Naval Postgraduate School*, 2020.

Khurshudyan, Isabelle, and Kostiantyn Khudov. “U.S.-Supplied Himars Changing the Calculus on Ukraine’s Front Lines.” *MSN*, The Washington Post , 1 July 2022, **https://www.msn.com/en-us/news/world/us-supplied-himars-changing-the-calculus-on-ukraine-s-front-lines/ar-AAZ5FnA**.

Lee, Timothy. “Autopilot Was Active When a Tesla Crashed into a Truck, Killing Driver.” *Ars Technica*, WIRED Media , 16 May 2019, **https://arstechnica.com/cars/2019/05/feds-autopilot-was-active-during-deadly-march-tesla-crash/#:~:text=A%20Tesla%20Model%203%20had%20Autopilot%20active%20in,divided%20four-lane%20highway%20in%20Palm%20Beach%20County%2C%20Florida**.

Li, Jian-hua. “Cyber Security Meets Artificial Intelligence: A Survey.” *Frontiers of Information Technology & Electronic Engineering*, vol. 19, no. 12, 10 Jan. 2019, pp. 1462–1474., **https://doi.org/10.1631/fitee.1800573**.

Liang, Percy, speaker. *Overview Artificial Intelligence Course | Stanford CS221: Learn AI (Autumn 2019)*, 6 Jan. 2020, **https://www.youtube.com/watch?v=J8Eh7RqggsU&t=547s**.

Lopez, Todd. “U.S.-Provided HIMARS Effective in Ukraine.” *U.S. Department of Defense*, DOD Newa, 15 July 2022, **https://www.defense.gov/News/News-Stories/Article/Article/3095394/us-provided-himars-effective-in-ukraine/source/us-provided-himars-effective-in-ukraine/#:~:text=U.S.-Provided%20HIMARS%20Effective%20in%20Ukraine%20July%2015%2C%202022,Russian%20forces%20who%2-0have%20illegally%20invaded%20that%20country**.

Luciano, Laoise, et al. “Digital Forensics in the Next Five Years.” *University of New Haven*, 30 Aug. 2018, https://doi.org/10.1145/3230833.3232813. Malewar, Amit. “Israel Used First-Ever AI-Guided Combat Drone Swarm in Gaza Attacks.” Inceptive Mind, 6 July 2021, **https://www.inceptivemind.com/israel-used-first-ever-ai-guided-combat-drone-swarm-gaza-attacks/19940/**.

McCarthy, J., et al. “A Proposal for the Dartmouth Summer Research Project on AI.” *Dartmouth AI Summer,* 31 Aug. 1955, **http://raysolomonoff.com/dartmouth/boxa/**.

Mount Vernon Ladies' Association of the Union (2022). "Spying and Espionage." *George Washington's Mount Vernon*. **https://www.mountvernon.org/george-washington/the-revolutionary-war/spying-and-espionage**. Last visited 27 September 2022.

Omni Magazine, 1987.

Pande, Jeetendra, and Ajay Prasad. Digital Forensics. Uttarakhand Open University, 2016.

Qadir, Abdalbasit Mohammed, and Asaf Varol. "The Role of Machine Learning in Digital Forensics." *2020 8th International Symposium on Digital Forensics and Security (ISDFS)*, June 2020, **https://doi.org/10.1109/isdfs49300.2020.9116298**. (H)

Ragsdale, Daniel. "AI Forensics; Defense Research and Engineering Interview ." 2022.

Sammons, John. *The Basics of Digital Forensics: The Primer for Getting Started in Digital Forensics*. Syngress, 2015.

Scharre, Paul, director. *Paul Scharre: Army of None: Autonomous Weapons and the Future of War*, Carnegie Council for Ethics in International Affairs, 1 May 2018, **https://www.youtube.com/watch?v=Z9hVOCUUBvM&t=2346s**.

Scherer, Matthew U. "Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies." *SSRN Electronic Journal*, 31 May 2015, **https://doi.org/10.2139/ssrn.2609777**.

Simon, Herbert A. *The New Science of Management Decision*. Harper & Row, 1960.

Steck, Harald, et al. "Deep Learning for Recommender Systems: A Netflix Case Study." *AI Magazine*, vol. 42, no. 3, 2021, pp. 7–18., **https://doi.org/10.1609/aimag.v42i3.18140**.

Tangredi, Sam J., and George Galdorisi. *Ai at War: How Big Data, Artificial Intelligence, and Machine Learning Are Changing Naval Warfar*e. Naval Institute Press, 2021.

Turek, Matt. "Explainable Artificial Intelligence (XAI)." *DARPA RSS*, Defense Advanced Research Projects Agency, 2018, **https://www.darpa.mil/program/explainable-artificial-intelligence.**

Walker, Richard. How AI Is Driving a Future of Autonomous Warfare | DW Analysis by DW News . *How AI Is Driving a Future of Autonomous Warfare*, Duetsche Welle, 25 June 2021, **https://www.youtube.com/watch?v=NpwHszy7bMk&t=232s**.

Wang, Yiran (14 February 2022). "Netflix's Recommendation Systems: Entertainment Made for You". Illumin Magazine, Univ. Southern California, CA. **https://illumin.usc.edu/netflixs-recommendation-systems-entertainment-made-for-you/**. Last visited 27 September 2022.

Yaacoub, Jean-Paul A., et al. "Advanced Digital Forensics and Anti-Digital Forensics for IOT Systems: Techniques, Limitations and Recommendations." *Internet of Things*, vol. 19, 23 May. 2022, p. 100544., **https://oi.org/10.1016/j.iot.2022.100544**.

Yampolskiy, Roman V. *Artificial Intelligence Safety and Cybersecurity: a Timeline of AI Failures*, 25 Oct. 2016, **https://doi.org/arXiv:1610.07997**.

Zhao, Rachel and Andra Valentina Krauze (April, 2021). Survival Prediction in Gliomas: Current State and Novel Approaches. In Debinski, W. *Gliomas* [Internet]. Exon Publications, Brisbane, AU. **https://www.ncbi.nlm.nih.gov/books/NBK570711/**

The Richard S. Schultz '60 Symposium Fellowship was established in 2017 as an endowed fund in honor of "Dick" by his wife of fifty years, Myrna L. Schultz, their children Marni and Alan, and his classmates and friends.

The fellowship enables Norwich undergraduates, from any academic discipline, the opportunity to pursue areas of inquiry and experiences that will promote and expand their understanding of the past and how it impacts the present and future. Through research, travel, and inquiry the Schultz Fellow and faculty advisor will offer perspectives for us to face the future with better understanding and confidence.

After a wide solicitation and competitive selection process, this annual fellowship is granted by the Norwich University Peace and War Center to an undergraduate student for a single project that may involve additional Norwich undergraduate students.

1819
NORWICH
UNIVERSITY®

Norwich.edu