Ben Walker
8-1-18

AI and National Security

Introduction:

AI and its use in security has been a continuing hot topic. From how it is used on defense, to how it can be used for offense, the discussions are quickly growing. However, in the little bit of time I have spent researching, I have found that within the umbrella of AI and its use in security, the discussion of AI and National Security is an extremely interesting and far reaching topic. It goes beyond if we should be concerned about if AI will be used to hack into our home computers or businesses, and instead applies to how governments and militaries may already be using AI to keep their country safe. I must note however, that the word safe, can have various meanings depending on who is using it. Such as much of my research has found, that just because one country thinks one way is the safest, other countries might believe otherwise. AI and its use by governments and militaries encompasses both the offensive and defensive sides. To be able to begin to grasp how AI is, and can be used, you must recognize that the ideologies of countries differ. In this way the nuances of what is thought of as right may be represented as contradictions in the usage of the technology, where sometimes it is used for defense, and sometimes offense. The call for better policies of how to handle AI in this way has already been made. For now, it is a bit of just a waiting game to see what will happen.

The Defensive Aspect:

Defense is one of the main focuses of AI and its use in security. Defensive AI, what it is, and the principals of how it is used, is one topic that scales to the national security level. Pattern prediction for subversive groups has direct application to national security. As well the ability to solve for a safe way to denuclearize the world, and the belief that AI should not be used for surveillance by the government and military are direct correlations to how AI can and must be considered for national security.

Defensive AI is one topic that many people say should happen, however what they mean isn't always so clear. From what I have found, when defensive AI is referenced, people tend to be discussing AI that can spot incorrect information. What I mean by this is anything that is designed to fool the AI, thereby allowing it to allow something to happen that it shouldn't[1]. This could be something like allowing a connection through a firewall or allowing someone to withdraw money when they shouldn't be. Many times, the issues that occur are due to the training of the AI being insufficient. A newer strategy for training AI makes us of GANs (Generative Adversarial Networks)[2]. This makes use of two AI, one to generate information, with the other to try and identify the bad information. In this way, the AI can learn how to more completely and correctly Identify the correct types of information and not throw false positives or false negatives. By training the AI like this, the capability of the AI at its task increases. This allows for them to be used as defensive AI where they can be relied upon to keep a system safe.

It is known that AI are very good with patterns, after all that is the basis of what an algorithm is. This allows for AI to be able to work with anything that has a pattern. One example that has been already put in place in China is using an AI to predict the formation of dangerous groups of people[3]. Now it must be noted that who China considers to be dangerous can take on multiple meanings, from those who disagree with the government, to those who wish citizens harm. This exact type of pattern prediction can also be applied to potential terrorist groups and allow us to predict what type of actions they might take. Currently we have humans doing this job, but this may allow for much of this work to be offloaded onto a machine. I would expect to see an AI like this, to eventually gain enough information to be able to make connections from past events, to point to how to prevent these things in the future, and who might be the ones to watch for[4]. Now this has been something to be concerned about as it could lead to too much control from the governing body. Yes, it is beneficial to protect from terrorists, but the aim is not to interfere with the citizens lives too much. This technology could be used to institute policies that might go against the rights of the citizens which would cause larger issues.

In a much more benign application, IBMs Watson, has already been used to help identify ways that it would be possible to denuclearize the world safely[3]. This is something that requires very complex problem solving. In my understanding, to do something like this, it would require the ability to account for enough variables that each country and its ideology would be looked at impartially. In my opinion, to truly get to the solution, it is necessary to analyze how the people interact, from the leaders of the countries, to the government within the countries. To be able to do something like that, would be an incredible accomplishment, and would be quite impressive if an AI was the source of the solution. This of course is a less disputed way of using

AI. Many people are concerned about AI being used to try to control, or influence for malicious intent, but this is something where the intent seems to be completely benign.

To continue off the concern over the usage of AI, many people believe that it should not be used for military or government surveillance[5]. This is something that would draw the line so that potentially the same type of system that would be used to track dangerous groups may be disallowed due to it being a form of surveillance. This would be an ideology that is designed with the best interest of citizens at heart. To keep them safe from a government or military that might seek to harm or control them. This too is a form of national security, that defends against internal issues. Since it is known that AI are good at data manipulation and sorting, that the scale of how much surveillance could be done could be well beyond what just humans could accomplish. After all, AI are very good at the task they are trained on and will be faster at it than a human[6].

The Offensive Aspect:

AI and their offensive side has many facets. From how people will seek to attack AI and try and break them, to how they will be used to attack and influence. AI are based off patterns, which can be mapped, as well they can accomplish jobs that have patterns. AI are ability to spear phish faster and more effective than humans can. AI are very good at data and information manipulation and may be able to easily influence people in this way through information warfare. Due to this control over antigovernment ideologies may be at risk, and the potential for a dystopian future with control over the populace. AIs use has been discussed and

possibly already planned for in military, for drone control, and missile guidance. The application of AI to create smart malware is also a possible risk to be aware of.

AI are based off patterns, they are ultimately an algorithm that integrates patterns into itself for how it should work. However, this allows for someone to be able to see how the AI operates and potentially figure out its weaknesses in this way. Generally, the method used is one where the attacker will try and feed information to the AI and then observe how it responds[1]. This can allow the attacker to map how the AI works if they are observant enough and know what they are expecting to have happen. One noted case of this is how someone might try and break through a AI meant to guard bank transactions from fraud. If the attacker can figure out the correct amount of money to try and steal, they can sneak under the radar of the AI and potentially get away with quite a lot of money; possibly through multiple small transactions[1]. This also applies to how an AI that might be monitoring a firewall could be tricked. If the attacker was to observe what type of traffic it was disallowing, they may be able to find one that it ignores and sneak through that way.

This problem is also related to how the attacker knowing that the AI is a pattern, may be able to train the AI after deployment, by feeding it information that it might mistake for good data and allow the data, thereby marking that data as always good. This is something where defensive AI are meant to be implemented to avoid, however the AI is only as good as its training. AI lack the creativity and abstraction necessary to be able to come up with very niche cases that a human may try. Even using something like GANs may be limited if the training AI is not able to generate enough cases to account for fringe cases. This points to the vulnerability of how AI function and that they are limited in their capacity completely to how good their

training is. This can allow larger more complex AI to more easily overwhelm and potentially control smaller AI. Since the larger AI have more cases they can use, they can try more things, potentially find those weaknesses in the smaller AI, and train them to accept certain information that causes them to behave as desired. Admittedly this idea dives into how AI warfare may work, and may not be something we see right away, but for national security, it is vital that it is accounted for. To ignore how AI warfare may happen and try to implement AI for national security will only lead to failing in the goal of keeping the nation safe.

Spear Phishing has been a topic of discussion with regard to AI, as an AI set up to spear phish will be able to act and react much faster than a human can, potentially allowing the effects to be much wider spread in less time[6]. This applies to an easy way of infiltrating systems, from getting malware onto systems, to gaining user information. I believe this could allow an AI to gather enough user data to be able to identify trends in user credentials even the ability to correctly guess what type of passwords may be used on a system. Maintaining a condensed database of usernames and passwords that it may be able to pick from in a more targeted fashion, may potentially allow AI to perform less of a brute force attack, and more of a probability-based attack. For the ability to install malware, while becoming less well used, can still give an easy in, especially to critical systems, and even make attacks like ransomware or snooping software more dangerous as now that there is a direct pipe. I will be covering more of how this will matter when I discuss smart malware.

Information manipulation and information warfare are concerns now, especially due to events such as what happened during the 2016 election in the United States. This was a showcase of how the bending of information shown could influence how people perceived

things. What the concern has been with regard to AI, is that it has the ability to be able to generate information and data such that it could create falsehoods that may be very believable[5]. The concern is especially relevant with how it could be possible to train an AI to edit photos, create matching text such that it could show a world leader doing something they never had, and spread the information around so that it appears legitimate. This has been noted as a possibility with dangerous consequences[5]. Such an attack while even being able to be proven wrong in the long term, may cause short term damage with long term consequences. An example of this is how information was able to be manipulated by Cambridge Analytica, such that they were able to claim that they swayed the election[7]. If something like this is able to happen at the speed of humans, the speed of machines will far exceed this potentially allowing for a type of spear phishing that can effect elections even down to local levels. While it is might not be thought of as a true phishing attack, the ability to get someone to click on a link or pay attention to something so that they give a desired response I would consider phishing. The major concern is how the ability for foreign entities, especially those with less funding to be able to sway larger entities in this way. The ability for countries to gain prowess regardless of military spending, will change the battlefields in many ways. The ability to sway other countries with AI must be considered a military action as cyber warfare is a very real battle. One such way that one such attack might succeed, is that a small country in Africa, wants to influence something in the United States. Now it is known that the Nigerian price style attacks tend to have errors in their English, and inconsistencies in how they work. AI level this playing field, as they can handle translation already. Many chat bots already exist as well. This type of technology could be applied and then have this AI carry on an attack such that many people

invest money into this country, thereby allowing the country to make back all the money it spent on the attack, and boost their country, all while the victims may be thinking they are donating to an aid cause, or even something on home soil. The way the information can be bent and manipulated does not stop there. With the rise of cryptocurrencies, and the refusal to outlaw them, it allows for untraceable transactions. This could allow money to disappear easily, and the chaos of it all to happen where no one could know who really began the attack. Could it have been someone in another country who began it and then gave the money to the smaller country to make it look like it was them? Or maybe it really was them? These will be the questions that will have their answers increasingly difficult to find. Information manipulation is a far reaching and dangerous attack as it could start up riots by showing false information, or get candidates elected or impeached as it suits another country. To be able to defend against this, having something like a defensive AI trained to spot posts or listings of this false information to swat it down before it gets published to the public. This however is a form of surveillance that many people may not be ok with. This could cause a chain reaction such that the information warfare attacks will not be blocked. The saving grace is that this could fall in line with both preventing the government and military from using AI for surveillance. Instead just mandating that the companies themselves must use surveillance to keep those on their service safe, no matter what country they are from. I will note that this argument hinges on the trust of government, and without that trust, the defense will not be able to work, yet at the same time, that trust is such that could allow the government to gain control and power over citizens that otherwise would be prevented.

Allowing governments to gain too much power and have too much control, is a topic that many people write about and are concerned about. Such concerns end with ideas of dystopian futures, or suppression at the hand of the government. One example of this is how China has an AI that they claim can allow them to know where dangerous crowds may form, or even give the citizens a form of dissention score[3]. This is something where understanding how the meaning of dangerous crowds, could be interpreted as those with ideologies that oppose the current government, may point to how these fears are justified. China is known for exerting this type of control over its citizens. Something where a government or military could be spying on the citizens it is supposed to be protecting, could allow for negative situations where those with ill intentions could use the information gathered to harm those who should be innocent. The ease of putting an AI to task in surveillance, can cause many jobs to be unnecessary, as well as allow for the potential risk of failure of the AI at doing its job correctly. To keep the citizens under surveillance is something that implies the citizens are guilty and can create further distrust of the government. The need of trust of government and trust by government is a balance that must be kept. Without this, the dystopian worlds will come true having citizens unable to create or innovate without concern of being labeled or marked. As well it potentially allows for the disruption of citizens being able to innovate something that may change how the world works for the better, but not necessarily in a way that benefits the government. The issue of what can potentially be done with the data that is gathered, is something that leaves too much availability for issue. From what happens if someone finds a way to gain access that should not have it, to what type of manipulation could happen if it becomes possible for someone to use the data to further an agenda. Maintaining the standpoint of preventing the

government or military from using AI for surveillance forces a form of democratization of the technology, where private companies would have to be the ones to use it in that method, and as such would be easier for the people to keep in check. While having the ability to spy on other countries may be beneficial to some, the issues that could be created if it is misused, seem to outweigh that benefit. This whole issue however may need to be treated much like nuclear weapons. The either everyone or no one must have this ability. As soon as someone has it, everyone must then play catch up[5]. The policies surrounding the use of AI must be established beforehand so that all governments can agree on them, possibly something like the Geneva convention even. In this way each country may be required to have a separate body verify that their AI are being used in the correct fashion.

The use of AI for the military is something that both interests and concerns many people. The ability for a government to be able to fund attacks that would use AI could be far less expensive than current technology. Something like drone control on a large scale could completely remove the need for human lives to be on the battlefield. I will note that there are UAVs that already implement AI for flight navigation such that they are point and click controlled. Northrop Grumman developed a UAV that could fly on its own and was working on teaching it to land on aircraft carriers, as well their most recent achievement being in air refueling[8]. This is something where we can already see that it is possible for AI to be used for control of individual aircraft. However, the consideration must be made for multiple aircraft, and the usage of AI for how they work together. This could potentially point to usage for low cost drones to be manufactured and equipped for tactical strikes or deliveries of payloads on massive scales. Thousands of drones could be used to work in sync with each other. A

demonstration of this is how Intel controlled drones for a light show used at concerts, as well as the Opening Ceremony for the Winter Olympics hosted in South Korea[9]. The concern of this technology being militarized is something that is real. There are rumors of a chip that contains many processors and contains an AI that is supposed to be equivalent to a human brain, that could fit inside a smart phone. Something like this could allow for AI to be put just about anywhere on a battlefield. From mobile ground units, to air units, the ability for them to be navigated without needing someone to do more than press an enter key is already here, and so all that is needed is the hardware to be deployed. Anywhere there is a risk of human lives being lost on the battlefield, there is the potential for an AI to be used instead. A command control AI could also be used to learn as it goes and make the battlefield function much like an AI in StarCraft 2 controlling its units. The look of the battlefield will change, and the need for any human lives to be risked may be negated. However, the importance of jamming technology in this situation will increase. Currently the United States Air Force is considering how it will use AI to be able to automatically control jamming frequencies which is listed as currently being manually done[3]. This jamming on the both the offensive and defensive sides being AI controlled will require each drone to need AI on its own to keep operating without direct command. This is much how battles are fought now. Again, we see there is a pattern in this, proving that AI can be used for it. Guided missiles are also another point where AI implementation is possible. This could allow for missiles to be able to be flown on paths and in ways that are currently impossible. Potentially having a missile that could fly at 20 ft above the ground avoiding radar detection, could allow for attacks to be done without needing to spoof radar systems like have been done in the past[10]. This could allow for an attack to be launched and that attackers clear

out before the attack is detected or even lands, with minimal effort. This would also be made possible as the need for direct sighting or targeting will not be needed. Identifying a convoy from a stealth UAV, not needing anything more than to know the general position of the convoy and being able to fire missiles that self-navigate and identify targets along with communicating between themselves of which target they are going after. This is an ability that makes military action much easier to complete, and potentially much less expensive as well. These kinds of benefits may cause countries to consider using this type of force rather than pursuing other options. This technology opens the gateway for dangerous possibilities to occur easily.

Malware itself has also been a concern especially when considering how AI could be applied to it. By developing malware with AI built in, you can end up with what I call smart malware. It can have the ability to change how it looks or operates, to disguise itself from antivirus or antimalware software. This can make malware particularly dangerous as it would now have the ability to disguise itself and possibly be undetectable with conventional methods. This could result in the ability to perform ransomware attacks on a targeted basis, or snoop through a computer and gather information about the user and what they do. Since AI are good at information manipulation and handling, once an AI is on a system, it could then sort through data on the machine looking for key things as well as be able to capture the browsing on the machine so that anything that could compromise the owner of the computer could potentially be sent out to be used elsewhere. This could also be particularly bad if machines that have access to sensitive information are attacked, as then there is a potential for that information to be stolen. As well it is known that infrastructure is very poorly guarded[11], so something like an

AI attacking and infecting computers connected to the infrastructure network, could potentially allow for it to gain control over whole grids and cause havoc to what is quickly becoming necessities of life for us. This kind of attack could easily be done by a foreign power. A somewhat similar attack that was carried out by Israel, when they spoofed the entire Syrian radar system, so that Syria was unaware anything was happening until the bombs and missiles began to land. This kind of attack, gaining control over systems like this, and becoming more automated with AI makes it all the more dangerous. This could also lead to the potential of a fully automated battlefield.

Closing thoughts:

The danger AI presents from how easy it is to use it for spying, or surveillance, is one that reaches beyond just how it could affect who can access a system, or make a transaction, but even what country has the power to sway another. In this way, AI can be seen that it is more dangerous than originally expected depending on how it is used. Yet it is so versatile that how it is used must be governed. AI can be implemented in so many ways that it could lead to a fully automated battlefield. The ability to wage war at the click of a button, making it more like a video game, and less about the lives at stake. The ability to completely remove humans from the battlefield and allow machines to control how attacks are done. A war waged at a low cost can allow smaller countries to potentially gain an upper hand as it does not matter so much about manpower anymore or how big a defense budget is. Everything from attack patterns, to jamming, to spoofing, to navigation, to command control systems could allow for AI to become

the ultimate weapons. This is why they must be handled and though about much like nuclear weapons. How they are used must be regulated and third parties must do checks to ensure the policies stated and agreed upon by the world are being upheld. Without the ability to manage the use of AI there will be an arms race as soon as it is realized that one party has an upper hand. Once one country finds a way to do something better, the rest must then catch up. This game of catch up will be forced to be played, as those who do not play, already lose. The ability for anyone to gain such a powerful weapon increases its danger level as there are many people who would seek to do harm. If we fail to correctly establish how to use AI and how to manage it, it will create a world where continuous war could be possible. It is possible that the first world power after AI warfare could be the only world power, as the AI that survives could be sufficient that it cannot lose. This type of world where the power balance could swing at the press of button is clearly dangerous to the continued survival of humanity. The danger present must not be ignored. Even though AI could be used for good and beneficial purposes, the use in war will develop it to win regardless, in this case, we humanity, ultimately lose.

References:

1) https://www.rtinsights.com/artificial-intelligence-gets-even-better-with-defensive-ai/
2) https://www.technologyreview.com/s/610656/to-protect-artificial-intelligence-from-attacks-show-it-fake-data/
3) https://www.roboticsbusinessreview.com/ai/ai_competition_seen_as_key_to_national_security/
4) https://www.binghamton.edu/news/story/440/researchers-can-predict-terrorist-behaviors-with-more-than-90-accuracy
5) https://thebulletin.org/2018/02/artificial-intelligence-and-national-security/
6) https://gizmodo.com/hackers-have-already-started-to-weaponize-artificial-in-1797688425
7) **Malice Domestic: The Cambridge Analytica Dystopia; Hal Berghel, University of Nevada, Las Vegas; Computer, www.computer.org/computer; May 2018**
8) http:www.nothropgrumman.com/Capabilities/X47BUCAS/
9) https://www.intel.com/content/www/us/en/technology-innovation/aerial-technology-light-show.html
10) https://www.wired.com/2007/10/how-israel-spoo/
11) https://www.nbcnews.com/news/us-news/u-s-infrastructure-can-be-hacked-google-simple-passwords-n548661