Information & Security

Published since 1998

G. Sharkov
vol. 49 (2021): 49-69
https://doi.org/10.11610/isij.4907

ISSN 0861-5160 (print), ISSN 1314-2119 (online)

Research Article

Harnessing the Potential of AI Against Covid-19 through the Lens of Cybersecurity: Challenges, Tools, and Techniques

George Sharkov [1,2] (✉)

[1] *European Software Institute – Center Eastern Europe, Sofia, Bulgaria https://esicenter.bg/*

[2] *Cybersecurity Laboratory at Sofia Tech Park, Bulgaria, https://sofiatech.bg/*

ABSTRACT:

Artificial Intelligence (AI) has already matured to the point where people from different industry fields and research domains can utilize its tools for various practical applications, including within healthcare, where AI holds immense promise. This capability has generated high hopes and has been used in the fight against the Covid-19 pandemic. However, against the backdrop of these applications, more contentious AI use cases have been highlighted, particularly with concerns about ethics and cybersecurity. This paper discusses recent developments and exciting applications of AI methods and tools. We cover three aspects: AI against Covid-19, AI for cybersecurity, and cybersecurity for AI, as well as some misuses of AI. We also present an overview of AI's potential through the lens of cybersecurity, to provide food for thought on the idea that securing artificial intelligence necessitates an entirely new approach to security and what it means in the context of dramatically increasing digital dependency.

ARTICLE INFO:

RECEIVED: 03 NOV 2021
REVISED: 29 NOV 2021
ONLINE: 05 DEC 2021

KEYWORDS:

Artificial Intelligence, AI, Cybersecurity, Covid-19, Ethical AI, Trustworthy AI

✉ Corresponding Author: George Sharkov; E-mail: gesha@esicenter.bg

Introduction

Achieving a machine-based imitation of human thinking and cognition may have looked a futuristic dream some decades ago, while nowadays, it is seriously penetrating our digitized ecosystem and we are really concerned about its possible and realistic impact on our privacy, freedom, and life in cyberspace, as we know it.

Although our progressively digitally-dependent society is still relatively far from answering Alan Turing's challenge from 1950 – "Can machines think?" the research and the practical implementations in the field of Artificial Intelligence (AI) have already reached a level of maturity, which makes the new old statement for AI as "the next big thing" this time better grounded. The third boom we presently observe, after the introduction of the term and the concept for "AI" by John McCarthy and colleagues in 1956 at the Dartmouth Conference and the following two "AI winters" in the 70s and 90s, is based on the recent developments suggesting a dramatic change on a large scale. Upgraded tools and techniques, such as Machine Learning (ML) and Deep Learning, Natural Language Processing (NLP), Computer Vision, Robotics, Autonomous Vehicles, and the reborn expert (intelligent, knowledge-based) systems have achieved widespread applications in industry, science, personal life, and society, including within the healthcare sector.[1]

Historically, namely biomedical research, healthcare, and medicine were among the promising areas during the second bloom of AI in the mid-1980s, when the MYCIN expert system demonstrated better-than-human performance in diagnosing the causes of blood diseases. Various similar healthcare-related expert systems followed, but only a few of them made it out of the research labs. Bulgaria has also demonstrated some applied research of AI-based systems and implementaiton, like in the field of medicine and biophysics (PREFES, KREBS [2] at the Institute of Biophysics, Bulgarian Academy of Sciences). Planning of biophysical experiments was facilitated by expert systems using fuzzy reasoning, and also observaitons such as microscope images and videos have been interpreted using qualitative-quantitative behavior analysis.[3] However, despite the promising results, some scientists have warned that in order to become a valuable decision-making adviser in real-life, these expert systems would need much better "common sense" knowledge formalized properly in addition to their experts' knowledge, plus a huge computational power to process large data sets. Those deficits were agreed as the main reasons for the second "AI winter" in the 1990s.

Nowadays, interest in AI for healthcare is back with much higher expectations and hopes, mainly based on the technical achievements providing huge computational power, access to massive volumes of data, and advanced machine learning methods. And it has a much better chance of succeeding on a large scale.

Recent AI developments allow people from different backgrounds to use its tools and techniques, such as Machine Learning (ML), Natural Language Processing (NLP), Computer Vision, and Autonomous and Expert/Reasoning Systems in many areas, including within the healthcare sector.[4] This potential has been largely mobilized in practice with the fight against the COVID-19 pandemic.[5]

Although often mentioned as somewhat of a buzzword, the topic of AI has been present in the media forefront, even more so than ever, since the onset of the SARS-CoV-2 virus pandemic. And not surprisingly so – within the growing body of research in the field of AI and its related models and technologies, serious concerns about data privacy and security are often raised.

Machine Learning instruments have been widely used to analyze existing data and produce predictions regarding the virus spread,[6] recognize patterns to help diagnose new cases [7,8] and help explain treatment outcomes.[9] Another promising trend of AI-based solutions to support the fight against the SARS-CoV-2 virus is for the rapid funneling of drug compounds, which could prove effective against the virus, and the analysis of potential side effects and medication effects predictions.

Despite the success of these applications, however, more controversial uses of AI have recently been reported, especially in the field of facial recognition, quarantine, and social distance control. As AI models need to process vast amounts of data, including sensitive and personal information, to be trained and integrated within a particular application, this makes them vulnerable to attacks from malicious actors and, in some cases, questionably unethical.

In this paper, we analyze and discuss some uses of artificial intelligence in the fight against the COVID-19 pandemic through the prism of cybersecurity, data privacy, and AI safety, including an overview of some challenges, opportunities, and tools. Furthermore, and to give more context, we will provide an overview of recent developments and exciting applications of AI methods and tools in terms of AI against COVID-19, AI for Cybersecurity, Cybersecurity for AI, and Misusing AI. Last but not least, we will doscuss that securing artificial intelligence will require new ways of thinking about security and what it means in the context of digital dependency and digitalized information society.

AI for Pandemic Control

Although AI-empowered applications have been reported as experimental used during various disease outbreaks, the unprecedented spread of the COVID-19 pandemic mobilized all means to tackle different aspects of disease spread and treatment.

The rapid adoption and early results have demonstrated that AI could play an essential role in the fight against COVID-19 and future disease outbreaks. Based on big data analytics, machine learning (ML), and deep learning (DL) methods and techniques, AI has proven helpful in identifying the disease's spread, clustering, trends, and patterns. They are also successful in predicting future outbreaks and mortality rates, supporting the diagnosis,[10,11,12] by monitoring large numbers of cases, resources and supplies management, and, of course, facilitating research for the prevention and effective treatment. Most of the applications evolved quickly from experimental to practical use, and even proof-of-concept pilots turned out to be of significant help in the fight against the "unknown."

AI for outbreak prediction

AI/ML was widely applied in systems to forecast the spread of the virus, produce early warnings, provide helpful information about the disease outbreak and vulnerable regions, and predict morbidity only by monitoring social media, blog posts, and news.

The Canadian startup company BlueDot has developed a system for early detection of the virus using AI. It is based on continuous review over 100 data sources of news, even ticket sales, climate data, demographics, and animal populations as well. They claim they detected the outbreak of pneumonia in Wuhan, China, on 31 December 2019 and that also could identify the cities most likely to experience this outbreak.[13]

Smart Wearables and Pre-Symptomatic Alerting

A new opportunity for AI-powered healthcare is what we call personal healthcare management. Personal healthcare management is made possible by the advent of new wearable technology, like smartwatches (e.g., Apple Watch), and activity/fitness trackers, such as Fitbit. These devices continually monitor features of our physiology, such as our heart rate and body temperature. This combination of features raises the fascinating prospect of having large numbers of people generating data series related to their state of health continually. AI systems could then analyze these data streams locally (via the more powerful smartphone you carry in your pocket) or by uploading them to an AI system on the internet, in the cloud.

It is vital not to underestimate the potential of these wearable technologies. For the first time, we could monitor our state of health continually. At the basic level, AI/ML-based healthcare systems may provide impartial advice on managing our health or generate an early alert. Some devices, like Fitbit, could not only monitor our activities but also set targets and monitor their achievement.

Coronavirus-infected people sometimes do not realize their symptoms for up to 5 days. In this situation, the virus can easily and asymptomatically spread to a large circle of people. A smart wearable ring manufactured by the Finnish startup Oura,[14] which records temperature, heart rate, respiratory rate, and activity levels, has been widely used for testing various AI-based algorithms for early COVID-19 infection alerting and reducing the spread of the virus. One of the models is claimed to predict whether the people infected within 24 h have COVID-19 symptoms and detect fever before one has it. The Oura ring can continually register various rest-taken schedules, action-based types and their degree, the ecosystem temperature, and pulse fluctuation in the body. The data collected from 65,000 subjects as part of the TemPredict study will be stored at the San Diego Supercomputer Center. They will be available to link with other datasets for further analyses. It is expected that AI/ML-based models analyzing the evolution and correlation of multiple parameters through data from wearable sensors will help early detect other infectious diseases, such as the flu.

Similar research at U.S. Army Medical Research and Development Command was piloted to monitor the health status of the personnel remotely and for AI-based early pre-symptomatic alerting. It is also expected that soon findings from this research will provide a near-continuous level of support and resilience to any U.S. Soldier across the globe.[15]

Remote Diagnosis and Telemedicine

The widest spread of AI/ML-empowered solutions for COVID-19 handling are the various telemedicine applications. By processing the data from remote sensors (like temperature, heart rate, respiration, and blood oxygen), such systems help clinicians care for patients in their homes, nursing homes, and hospitals and optimize triage and resource planning. Based on such parameters and other remotely detectable health signs, such as voice, movement, weight, and even toileting, the ML models are used to monitor and predict the onset of adverse events and the progression of the disease. AI was used to augment and adapt mobile health applications which use smart devices like watches, mobile phones, cameras, and other wearables for diagnosis, efficient monitoring, and contact tracing.

A non-traditional AI/ML-based method for early COVID-19 diagnosis was introduced by the MIT research team.[16] It uses a novel approach for early diagnosis of COVID-19 by analyzing the coughing as recorded and transmitted by a dedicated mobile application. The researchers have found that asymptomatic people differ from healthy individuals in how they cough, although this is not decipherable to the human ear. The model is trained on tens of thousands of samples of coughs, as well as spoken words. It accurately identified 98.5 percent of coughs from people who were confirmed to have COVID-19, including 100 percent of coughs from people without symptoms who tested positive. A cloud-based application is ready for production and further training on data sets (initially trained on more than 200,000 forced-cough audio samples). In previous research, ML-based methods have been used to detect defects in human vocal cords by pronouncing different phrases, such as “mmm” or “them,” and for the early diagnosis of Alzheimer’s disease by analyzing emotional states in speech.

AI for Monitoring Cases and Logistics

AI techniques are applied for monitoring patients in clinical settings and predicting the course of treatment. Based on the data derived from vital statistics and clinical parameters, AI was helpful in allocating resources and prioritizing the need for equipment, such as ventilators and respiratory support in intensive care units. AI can also be used to predict the chances of recovery or mortality in COVID-19, provide daily updates, storage, and trend analysis, and chart the course of treatment.

Researchers from Israel reported an AI model which predicts the length of COVID-19 hospitalization. They used AI/ML to track hospitalized COVID-19 patients between clinical states and predict the number of days expected in different states through a personalized model. The system was trained and validated through substantial data sets from the Ministry of Health and the COVID-19 hospitalized patient registry, which includes patient age and sex in addition to daily clinical status and dates of admission and discharge.[17]

Accelerating Research for Drugs and Vaccines

Because of the unpredictable yet highly contagious nature of the COVID-19 virus, research for analyzing the structure of the virus to create drugs and effective vaccines became the highest priority. The research is challenging since the virus belongs to a family of enveloped coronaviruses that contain single-strand RNA structures. Yet, similarly to double-stranded viruses, such as HIV, Ebola, and others, COVID-19 can rapidly mutate, making vaccine development and virus analysis difficult. AI methods and tools are being used to support this research and accelerate vaccine development.

A successful implementation of the Linearfold algorithm, disclosed by Baidu to researchers, is significantly faster than traditional RNA folding algorithms at predicting a virus's secondary RNA structure. Baidu AI scientists have used this algorithm to predict the secondary structure prediction for the COVID-19 RNA sequence, reducing overall analysis time from 55 minutes to 27 seconds, meaning it is 120 times faster.

The MIT researchers have used machine learning to identify medications that may be repurposed to fight COVID-19.[18] They have developed appropriate cell culture models to validate the hypothesis for a correlation between viral infection/replication and tissue aging and allow for highly specific and targeted drug discovery programs.

To help researchers generate potential new drug candidates for COVID-19, IBM has applied the novel AI generative frameworks to three COVID-19 targets and has generated 3000 novel molecules, shared with scientists. The researchers at the Quebec institute Mila have used ML to discover antiviral drugs to fight COVID-19, using graph neural networks to explore combinations of existing drugs and trying to search for all possible drug-like molecules.

AI/ML is helping in the race for the development of a vaccine against the pathogen.[19] Researchers from the University of Michigan [20] used their Vaxign reverse vaccinology-machine learning platform that relied on supervised classification models to predict possible vaccine candidates for COVID-19. Thus, AI has accelerated manifold the pace of discovery. The rapid development of two highly effective mRNA vaccines (from Moderna and Pfizer) was possible through AI technology and innovative collaboration among researchers worldwide.[21] Thanks to AlphaFold2, the AI system created by the London-based company DeepMind, it was possible to predict the three-dimensional structures of very challenging target proteins with high accuracy. It is also used to model the possible mutations of the virus and, thus, the improvement of the vaccines.

AI helps not only the discovery and evolution of vaccines. Moderna and IBM plan to use modern technologies, AI and blockchain, for smarter vaccination management, distribution, and supply chain management.

Chatbots and Service Robots

Popular chatbots have been quickly adapted and widely used to disseminate information, especially in remote settings, as well as for symptom monitoring, behavior change alerting, mental health support, and remote assistance. However, researchers and authorities have warned about important challenges.

Providing reliable and evidence-based information is critical in a pandemic, but there could be issues, such as conflicting advice between global and local authorities and misinformation. The developers should decide how to amplify reliable sources and coordinate global information sources, such as the WHO (World Health Organization), with advice from regional authorities.

Some AI-empowered web and mobile-based chatbots, such as the World Health Organization's Health Alerts or the Center for Disease Control's Covid-19 portal, are not only providing up-to-date information but also conducting self-diagnosis for coronavirus infection at home. Those portals were also emulated in many developing countries.[22]

Chatbots usually provide links to third-party services through which personal data might be shared with unexpected consequences. Symptoms screening and sharing health-related information between companies and governments are among the sensitive areas of application of these technologies. Besides, there is a boom of service and anthropomorphic robots with an AI core that can be used to deliver essential services and routine assistance.

Reforming Data Analytics

The pandemic has changed the way AI was traditionally used for data analytics. Previous ML-based applications have been widely used for big data analytics, including Deep Learning techniques. According to Gartner,[23] when COVID-19 hit, organizations using traditional analytics techniques that rely heavily on large amounts of historical data realized that many of these models are no longer relevant and a lot of historical data sets are useless. The forward-looking data and analytics teams are pivoting from traditional AI techniques relying on "big" data to analytics that requires less or "small" and more varied data and applying adaptive machine learning. In addition to the expected technology scalability, responsible and ethical AI norms should be implemented to avoid data bias and provide data privacy.

Examples and Highlights

- *AI and control of COVID-19 article by the Council of Europe.* "Artificial intelligence (AI) is being used as a tool to support the fight against the viral pandemic that has affected the entire world since the beginning of 2020. The press and the scientific community are echoing the high hopes that data science and AI can be used to confront the coronavirus and "fill in the blanks" still left by science."[24]

- *In Search for Cure: How Baidu is bringing AI to the fight against coronavirus.* In partnership with the Oregon State University and the University of Rochester, the Chinese company Baidu is intensively working on the Linearfold prediction algorithm. The algorithm studies the structure of the virus's secondary RNA, thus aiming to provide scientists with further information about how the virus is spreading, along with its evolutionary patterns.[25]
- *AI for Computational predictions of protein structures associated with COVID-19.* DeepMind continues improving the AlphaFold system while releasing its structure predictions of several under-studied proteins associated with SARS-CoV-2. Their experiments have so far confirmed aspects of their model, raising hopes about the possibility of drawing biologically relevant conclusions from blind predictions of even very difficult proteins and thereby deepening our understanding of understudied biological systems.[26]
- *IBM, Amazon, Google, and Microsoft partner with the White House to provide computing resources for COVID-19 research.* AWS has already dedicated $20 million to support COVID-19 research. Microsoft has announced several initiatives, mostly helping businesses cope with the fallout of this crisis.[27]
- *Predicting the evolution of the virus: The BlueDot case.* The Canadian company BlueDot is predicting the virus evolution thanks to its AI/ML-based algorithm, which looks at more than 100 datasets—including news sources, airline ticket sales, demographic data, climate data, and animal populations—to predict and track the spread of disease.[28]
- *Using AI to verify compliance with the anti-epidemic measures by phone.* AI has been widely used in support of such mass surveillance policies, with devices being used to measure temperature and recognize individuals or to equip law enforcement agencies with "smart" helmets capable of flagging individuals with high body temperature.[29]

AI for Cybersecurity: Examples from the COVID-19 Pandemic

The year of the COVID-19 pandemic will undoubtedly be remembered as the year in which cybersecurity events exploded, and cyber incidents transformed the way we live and work. Due to the intensified use of the internet and virtualization, cyber incidents have also increased dramatically. More than 445 million cyberattacks were reported in 2020, double in comparison with 2019.[30] But not only has the number and intensity increased but also the scope and the sophistication of the attacks have noticeably evolved, as the impact, as well as the motivation and the tools of malicious actors. Since the onset of the pandemic, the FBI has seen a fourfold increase in cybersecurity complaints and global losses from cybercrime estimated at above $1 trillion in 2020.[31] Several headlines already qualified 2020 as *"The year that the COVID-19 crisis brought a cyber pandemic."*

Of highest interest were all types of knowledge, data, and information related to COVID-19 research, drugs and vaccines, test results, and healthcare and patients' records in particular. In July 2020, the UK National Cyber Security Centre (NCSC) reported that drug firms and research labs had been targeted for Covid-19 vaccine information by a group known as APT29 (Russian state-sponsored hackers).

AI/ML methods and tools are already widely used in incident detection and prevention systems (IPDS), and more sophisticated and advanced SIEMs (Security Information and Event Management systems) for network and systems behavior monitoring, filtering "false positives," and rapid response. Due to the increased intensity of the attacks and the growing attack surface complexity, AI/ML methods and tools have become inevitable for threat assessment, effective cyber defense, threats assessment, and resilience.

The main types of growing attacks in 2020 and some of the novel AI-based threat-hunting methods and tools were:

- *Social engineering* - a third of the breaches, of which 90% by phishing - AI is used to detect various AI-enabled attacks, like "deep fakes" (technology can determine when an image or video is counterfeit). AI/ML is used to filter out fake reviews in a dataset (e.g., statistics show 61% of electronic reviews on Amazon are fake) and misinformation.
- *Ransomware* - just the ransom demands amounted to $1.4 billion, 22% of the cases. In Germany, cybercriminals targeted a hospital for ransom, with patient care systems being disabled, resulting in one patient's death.
- *DDoS attacks* remain a growing threat, with 4.83 million DDoS attacks attempted in the first half of 2020 alone. Since criminals now employ AI to perform DDoS attacks, AI/ML and behavior monitoring tools are the cure to look for the weak spots, especially if a massive amount of data is involved.
- *Supply chain* – third-party software, supply chain, and corporate security challenges – AI/ML/DL technology is for "hidden threats" analysis and remote working environment.

Examples and Highlights

- *Sensitive Content Filtering with AI: Facebook is now using AI to sort content for quicker moderation*. Facebook has made yet another step in the direction of having artificial intelligence handle more moderation duties on its platforms. Lately, it announced its latest step toward that goal: putting machine learning in charge of its moderation queue and limiting the need for human review of posts that include everything from spam to hate speech.[32]

- *Check Point Presents the First Autonomous Threat Prevention System*. Check Point, a global leading provider of cybersecurity solutions, has introduced its next-generation unified cyber security platform. The platform delivers autonomous threat prevention designed for the entire distributed enterprise.[33]
- *Machine Learning: Higher Performance Analytics for Lower False Positives*. Faced with mounting compliance costs and regulatory pressures, financial institutions are rapidly adopting Artificial Intelligence (AI) solutions, including machine learning and robotic process automation (RPA) to combat sophisticated and evolving financial crimes.[34]
- *Applicability of machine learning in spam and phishing email filtering: review and approaches*. Machine learning models are being extensively used by leading internet service providers like Yahoo, Gmail, and Outlook, to filter and classify UBEs successfully.[35]
- *A Machine Learning Study on Phishing URL Detection*. When the goal is to flag a suspicious phishing URL previously unknown to blacklist data providers, Machine learning offers a solution used for such a prediction task.[36]
- *Thorough Analysis For Using Data Science To Detect Malicious Domains*. Analyzing existing enterprise traffic logs with a data science approach is an efficient way to detect signs of a breach. VPN and Active Directory logs can be used to detect compromised account activities. Database or file-level access logs can also be used to detect insider threat activities. Mining these voluminous logs require different machine learning and data mining methods will vary depending on use cases.[37]

Towards Robust and Trustworthy AI: Cybersecurity for Artificial Intelligence

Undoubtedly, AI systems based on software and IT need to comply with evolving cybersecurity requirements. But this is not enough, as the AI methods and tools are based on different architectures, technologies, algorithms, and data compared to traditional systems. The EU approach to AI, as outlined in the EU Strategy for AI from 2018, and in the EC White Paper of February 2020 on AI is defined as "ethical, secure and cutting-edge AI made in Europe."

The Executive Director of the EU Agency for Cybersecurity ENISA Juhan Lepassaar said: "Cybersecurity is one of the bases of trustworthy solutions for Artificial Intelligence. A common understanding of AI cybersecurity threats will be key to Europe's widespread deployment and acceptance of AI systems and applications."

In the ENISA "AI Cybersecurity Challenges" report of December 2020,[38] an AI cybersecurity ecosystem and a "Threat Landscape for AI" are outlined. It is stated that "When considering security in the context of AI, one needs to be aware that AI techniques and systems making use of AI may lead to unexpected outcomes and may be tampered with to manipulate the expected outcomes. This is particularly the case when developing AI software that is often based on fully black-box models, or it may even be used with malicious intentions, e.g. AI as a means to augment cybercrime and facilitate attacks by malicious adversaries." It is, therefore, essential to secure the AI itself.

The steps to achieve cybersecurity for AI, specifically tailored for machine learning-based models and the AI development and implementation lifecycle, are:

- understand what needs to be secured (assets, subject to AI-specific threats and adversaries)
- understand the related data governance
- manage threats in a multi-party ecosystem in a comprehensive way by using shared models and taxonomies
- develop specific controls to ensure that AI itself is secure.

The cybersecurity threats to AI are listed as: "lack of robustness and the vulnerabilities of AI models and algorithms, e.g. adversarial model inference and manipulation, attacks against AI-powered cyber-physical systems, manipulation of data used in AI systems, exploitation of computing infrastructure used to power AI systems' functionalities, data poisoning, environment variations which cause variations in the intrinsic nature of the data, credible and reliable training datasets, algorithmic validation/verification (including the integrity of the software supply chain), validation of training and performance evaluation processes, credible and reliable feature identification, data protection/privacy in the context of AI systems, etc."

Cybersecurity is fundamental for trustworthy AI solutions, but three general aspects are listed in the "Ethics Guidelines for Trustworthy AI"[39] by the EU High-Level Expert Group:

- lawful
- ethical
- robust (with technical robustness and safety, security and resilience, transparency, traceability, explainability, etc.).

Standards and certification schemes are under development by standardization bodies (ETSI, CEN, ISO/ICE, and others). An "Assessment List for Trustworthy Artificial Intelligence" (ALTAI)[40] is available online for self-assessment, specifically tailored for SMEs.

Examples and Highlights

- *Microsoft announces two AI-based technologies to combat disinformation*. Microsoft announces two AI-based technologies for media analysis to detect manipulated content and assure the authenticity of a given media artifact. One of the solutions offers a browser extension to check certificates and match hashes, letting people know about the degree of accuracy and authenticity of the viewed content.[41]

- *Attacking Artificial Intelligence: AI's Security Vulnerability and What Policy-makers Can Do About It*. As AI systems are integrated into critical commercial and military applications, attacks against AI can have serious, even life-and-death, consequences. AI attacks can be used in several ways to achieve a malicious end goal. This report provides guidelines, general information, and recommendations to policy-makers about securing AI.[42]
- *Scalable Private Learning with PATE*. Federated Learning to secure Machine Learning Against Privacy Attacks.[43]
- *Employing Encryption Techniques on Machine Learning Training Data*. This paper presents a method to convert learned neural networks to CryptoNets, which can be applied to networks.[44]
- *Early Warning System for Disinformation Developed with AI*. Researchers at the University of Notre Dame are working on a project to combat disinformation online, including media campaigns to incite violence, sow discord, and meddle in democratic elections.[45]
- *Scientists voice concerns and call for transparency and reproducibility in AI research*. Scientists challenge scientific journals to hold computational researchers to higher standards of transparency and call for their colleagues to share their code, models, and computational environments in publications.[46]
- *Security software for autonomous vehicles*. Before autonomous vehicles participate in road traffic, they must demonstrate conclusively that they do not pose a danger to others. New software prevents accidents by predicting different variants of a traffic situation every millisecond.[47]
- *Making AI Trustworthy*. A new tool generates automatic indicators if data and predictions generated by AI algorithms are trustworthy.[48]

Misusing AI: Good Technology Gone Bad

Stephen Hawking famously said, *"AI will be either the best or the worst thing ever to happen to humanity."* Now, more than ever, and especially in the light of a global pandemic, we realize the double-edged sword that AI is if misused. Privacy concerns, AI algorithms tracking our every move, have come to the media forefront, and weaponizing AI has become increasingly scarier.

From using AI to remotely execute intelligent, self-propagating attacks to employing AI to track abidance to pandemic countermeasures or using ML to mimic the behavior of trusted system components, we have witnessed a lot of artificial intelligence misuse during 2020. And we are convinced now, more than ever, that AI will be either the best or the worst thing to ever happen to humanity.

Intelligent Surveillance

With recent developments in AI for video and audio analytics, the nature of what we think surveillance is becoming subject to change. Experts worry that besides some positive outcomes, such as AI-powered cameras being able to recognize people breaking the law or posing an immediate danger to others, troubling predictions are also becoming a reality. With powerful algorithms able to quickly identify people, this data can be further correlated to other data about the same person, providing a very indiscrete insight into people's lives, motivations, and behavioral patterns. Furthermore, with the increasingly cheaper and accessible cloud and hardware storage, video, audio, and other artifacts of our every move are being stored for longer than they used to be, making it easier to "dig up dirt," for instance.

Facial and Voice Recognition

With hundreds of bots automatically scraping the web for video and audio recordings, along with images of people, an enormous amount of data is being processed and analyzed without people's consent, creating vast facial or voice recognition databases for training a large variety of machine learning algorithms. This non-consensual collection of personal and sensitive data could put an end to privacy by falling into malicious hands or being used for questionable purposes. Besides, with the advancement of deep fakes, seeing no more equals believing, and we become increasingly troubled when attempting to recognize fake news, footage, recordings, and information.

Faking Medical Data and Images

Deep Learning malware samples have been explicitly tested in medical environments, showcasing a variety of intelligent attacks against images, such as altering MRI scans, or even more scary, altering a patient's diagnosis by recognizing and removing tumors from MRI scans. The possibility to seamlessly conduct intelligent AI-based attacks on entire systems-of-systems by intelligently mimicking components of a healthcare service or supply chain has also scared healthcare providers and cybersecurity experts alike.

Examples and Highlights

- *Deep Learning Malware Can Fake Cancer on Medical Images.* A deep learning algorithm successfully penetrated a healthcare organization and fooled both humans and an AI system with faked medical images. The algorithm infiltrates a typical health system's PACS infrastructure and alters MRI or CT scan images using malware based on a type of machine learning called generative adversarial networks (GANs) to inject fake tumors or remove real cancers from the patient data.[49]
- *Deepfakes Are Going To Wreak Havoc On Society.* We Are Not Prepared. The amount of deepfake content online is growing at a rapid rate. At the beginning of 2019, there were 7,964 deepfake videos online, according to a report from startup Deeptrace; just nine months later, that figure had jumped to 14,678. Even more troubling, it is certain that deepfakes will make it increasingly difficult for the public to distinguish between what is

real and what is fake, a situation that malicious actors will inevitably exploit.[50]

- *A surveillance company harassed female employees using its facial recognition technology.* A surveillance startup in Silicon Valley is being accused of sexism and discrimination after a sales director used the company's facial recognition system to harass female workers. Last year, the sales director accessed these cameras to take photos of female workers, then posted them in a Slack channel called #RawVerkadawgz alongside sexually explicit jokes.[51]
- *AI Has Made Video Surveillance Automated and Terrifying.* AI can flag people based on their clothing or behavior, identify their emotions, and find people who are acting "unusual."[52]
- *Clearview AI stops facial recognition sales in Canada amid privacy investigation.* Clearview AI will no longer sell its facial recognition software in Canada, according to government privacy officials investigating the company. The end of Clearview AI operations in Canada will also mean the end of the company's contract with the Royal Canadian Mounted Police.[53]
- *Protecting smart machines from smart attacks.* In a series of recent papers, a research team has explored how adversarial tactics applied to artificial intelligence (AI) could, for instance, trick a traffic-efficiency system into causing gridlock or manipulate a health-related AI application to reveal patients' private medical history.[54]
- *Security Attacks Analysis of Machine Learning Models.* An overview of common security risks and attacks related to ML.[55]

AI for Recovery and Development in the Post-Covid-19 Era

Finally, observing the already passing climax of the pandemic and the waning consequences, we began to plan the economy's recovery and restore our normal life. What are the lessons learned, and how could AI help not only in restoring but also improving our digitalized society? These opportunities have been outlined in the latest World Bank report,[56] "Harnessing Artificial Intelligence for Development on the Post-COVID-19 Era." The report summarizes the new and innovative AI applications and solutions to help manage the spread of the virus, drive drug discovery and cope with social distancing requirements.[57] A broader application of AI, at the onset of the pandemic, was used to monitor the spread of the virus and predict where and when new outbreaks might occur. AI was widely used to accelerate the speed of vaccine research, yielding multiple vaccines in record time. AI was also used to power web and mobile-based chatbots to help people find information quickly and conduct self-diagnosis for coronavirus infection at home. And as the pandemic enters the vaccination phase, some AI-based applications are applied to inform triage planning for population groups, to forecast demands, to manage supplies and supply chains and to monitor postvaccination of drug reactions.[58]

Based on the experience and lessons learned, there are three immediate areas where AI/ML could be of further help, namely:

- AI could help to limit the further spread of infection and continuously monitor for eventual new waves;
- AI could build on the lessons learned from the pandemic to alert earlier governments for eventual future disease spread and pandemic;
- AI may offer commercial benefits and facilitate economic and services growth by accelerating the technology revolution (Industry 4.0 towards 5.0).

The World Bank report [59] underlines that AI technologies' development remains heavily concentrated in a small number of advanced economies. This has immediate implications for global recovery efforts in the post-Covid-19 era, considering AI's potential to help developing countries rebuild quickly across critical sectors once the pandemic subsides. They have studied the leading countries' national strategies and approaches, ranking high in global indexes such as Stanford Global AI Index and the Oxford Insights Government AI Readiness Index. The purpose was to advise and guide the lagging countries. In 2020, Bulgaria accepted a concept for the development of Artificial Intelligence in Bulgaria until 2030, focusing on scientific advancements and software/IT development, and two priority sectors - intelligent agriculture and healthcare.

Conclusions

The "new normal" imposed by the COVID-19 pandemic during 2020 has tremendously accelerated the deployment of experimental technologies in healthcare, as well as and facilitation and security for virtualized teaming and teleworking. Likewise, a significant boost was given to the adoption of AI following the explosive growth of investments during the past decade.

Notwithstanding, the lockdown also gave a boost to the actual digital transformation of business, public administration, education, and social life and has forcedly activated plans belated for years. Although mostly experimental, AI developments have quickly been harnessed in the fight against the pandemic. Undoubtedly, this "third AI" boom will have a significant impact not only on healthcare but also on the completely new organization and quality of life in this "digitized to survive" world in which we currently live.

Artificial intelligence has become an important cornerstone of the fight against the COVID-19 pandemic, as well as in more and more areas of life, work, and leisure, but also an indispensable element of the future of the economy not only in the European digital market but worldwide alike.

The examples and overview provided in this paper aimed to shed light on some of the most recent developments and applications of AI through the prism of cybersecurity, using COVID-19 as a leveraging mechanism for innovation. Through this overview, we provided yet another example of the need for ethical, technically robust, and lawful AI, showcasing not only the challenging areas but also the achievements and opportunities opened to representatives from different industries using AI. This overview further provided a stage for the intersection between cybersecurity and AI to shine through as a necessary component to innovation and securing the future of freedom and fundamental human rights.

Finally, we put forth the proposition that potentially regulating Artificial Intelligence to ensure the development of secure, trustworthy, and technically robust AI solutions is a fundamental approach toward the evolution of the potential of AI solutions but also for protecting the essential human rights of citizens worldwide.

Acknowledgments

This research is partially supported by the British Embassy in Sofia through their Cybersecurity Newsletter initiative, which served as a baseline for this text.

References

[1] Sungho Sim and Myeongyun Cho, "Convergence model of AI and IoT for virus disease control system," *Personal and Ubiquitous Computing* (2021), https://doi.org/10.1007/s00779-021-01577-6.

[2] George Sharkov and D.S. Dimitrov, "Knowledge-Based Systems in Biophysics: Applications to Research in Neurobiology of Aging and Medicine," in *Artificial Intelligence III (AIMSA)*, edited by T. O'Shea and V. Sgurev (Elsevier Science Publishers B.V., 1988), 397-404.

[3] Boiko Balev and George Sharkov, "Knowledge-Based Interpretation of Biophysical Images," in *Artificial Intelligence IV – AIMSA*, editeb by Ph. Jorrand and V. Sgurev (Elsevier, North-Holland, 1990), 405-414, https://doi.org/10.1016/B978-0-444-88771-9.50049-0.

[4] Sim, "Convergence model," (2021).

[5] Yoshiki Oshida, *Artificial Intelligence for Medicine: People, Society, Pharmaceuticals, and Medical Materials* (Berlin, Boston: De Gruyter, 2021), https://doi.org/10.1515/9783110717853.

[6] Matissa Hollister, "COVID-19: AI can help but the right human input is key," World Economic Forum (March 2020), https://www.weforum.org/agenda/2020/03/covid-19-crisis-artificial-intelligence-creativity/.

[7] Wim Naude, "Articial intelligence against covid-19: an early review," *Medium* (April 2020), https://towardsdatascience.com/articial-intelligence-against-covid-19-an-early-review-92a8360edab.

[8] Qin C., Šídek A., Nelson AWR, Bridgland A., Penedones H., Petersen S., Simonyan K., Crossan S., Kohli P., Jones DT, Silver D., Kavukcuoglu K., Hassabis D., "Improved protein structure prediction using potentials from deep learning," *Nature* 577, no. 7792 (2020): 706–710.

[9] Markus Schmitt, “How to fight COVID-19 with machine learning towards data science,” *Medium*, April 2020, https://towardsdatascience.com/ght-covid-19-with-machine-learning-1d1106192d84.

[10] Prabira K. Sethy , Santi Behera, Pradyumna Ratha, and Preesat Biswas, “Detection of coronavirus (COVID-19) disease based on deep features and support vector machine,” *International Journal of Mathematical, Engineering and Management Sciences* 5, no. 4 (2020): 643-651, https://doi.org/10.33889/IJMEMS.2020.5.4.052.

[11] Apostolopoulos, I. D. and T. A. Mpesiana, “COVID-19: Automatic detection from X-ray images utilizing transfer learning with convolutional neural networks,” *Physical and Engineering Sciences in Medicine* 43, (2020): 635–640.

[12] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim et al., “Automated detection of COVID-19 cases using deep neural networks with X-ray images,” *Computers in Biology and Medicine* 121, no. 103792 (2020).

[13] Cory Stieg, “How this Canadian start-up spotted coronavirus before everyone else knew about it,” *CNBC*, Mar 6 2020, https://www.cnbc.com/2020/03/03/bluedot-used-artificial-intelligence-to-predict-coronavirus-spread.html.

[14] “The most accurate smart ring. Health tracking wrapped around your finger — track your sleep, activity, recovery in style,” *Oura Ring*, 2022 ,https://ouraring.com.

[15] Ramin Khalili, “For the Pandemic and Beyond, Wearable Technology Points the Way,” *U.S.ARMY*, January 13, 2021, https://www.army.mil/article/242364/for_the_pandemic_and_beyond_wearable_technology_points_the_way.

[16] Jennifer Chu, “Artificial intelligence model detects asymptomatic Covid-19 infections through cellphone-recorded coughs. Results might provide a convenient screening tool for people who may not suspect they are infected,” *MIT News Office*, October 29, 2020, https://news.mit.edu/2020/covid-19-cough-cellphone-detection-1029.

[17] Kat Jercich, “New AI model can predict length of COVID-19 hospitalization. By using patient age, sex and daily clinical state, the machine learning model also predicts the probability of in-hospital mortality,” *Healthcare IT News*, January 22, 2021, https://www.healthcareitnews.com/news/new-ai-model-can-predict-length-covid-19-hospitalization.

[18] Kat Jercich, “MIT researchers use AI to find drugs that could be repurposed for COVID-19. The research team noted that lung tissue gets stiffer as a person gets older, showing different patterns of gene expression than in younger people,” *Healthcare IT News*, February 15, 2021, https://www.healthcareitnews.com/news/mit-researchers-use-ai-find-drugs-could-be-repurposed-covid-19.

[19] Keshavarzi Arshadi A., Webb J., Salem M., Cruz E., Calad-Thomson S., Ghadirian N., Collins J., Diez-Cecilia E., Kelly B., Goodarzi H., and Yuan JS, “Artificial Intelligence for COVID-19 Drug Discovery and Vaccine Development,” *Frontiers in Artificial Intelligence* 18 (2020), https://doi.org/10.3389/frai.2020.00065.

[20] Edison Ong, Mei U Wong, Anthony Huffman, and Yongqun He, “COVID-19 Coronavirus Vaccine Design Using Reverse Vaccinology and Machine Learning,” *Frontiers in*

Immunology, 03 July 2020, https://www.frontiersin.org/articles/10.3389/fimmu.2020.01581/full.

[21] Sara Ibrahim, "The rapid development of two highly effective Covid-19 vaccines was made possible through AI technology and innovative collaboration among researchers around the world, including Switzerland," *SWI swissinfo*, January 8, 2021, https://www.swissinfo.ch/eng/artificial-intelligence-helps-bring-about-record-fast-vaccines/46256752.

[22] Adam S. Miner, Liliana Laranjo, and A. Baki Kocaballi, "Chatbots in the fight against the COVID-19 pandemic," *npj Digital Medicine* 3, no. 65 (Nature, 2020), https://doi.org/10.1038/s41746-020-0280-0.

[23] Kasey Panetta, "Gartner Top 10 Data and Analytics Trends for 2021," *Gartner*, March 15, 2021, https://www.gartner.com/smarterwithgartner/gartner-top-10-data-and-analytics-trends-for-2021/.

[24] "Artificial intelligence and the control of COVID-19. AI and control of COVID-19," Council of Europe*,* 2021, https://www.coe.int/en/web/artificial-intelligence/ai-covid19.

[25] Baidu, "How Baidu is bringing AI to the fight against coronavirus. Scientific and medical communities worldwide are using AI to understand and contain Covid-19 treat infected patients, and ultimately develop vaccines that prevent future outbreaks," *MIT Technology Review*, March 11, 2020, www.technologyreview.com/2020/03/11/905366/how-baidu-is-bringing-ai-to-the-fight-against-coronavirus/.

[26] "Computational predictions of protein structures associated with COVID-19," *DeepMind*, August 4, 2020, https://deepmind.com/research/open-source/computational-predictions-of-protein-structures-associated-with-COVID-19.

[27] Frederic Lardinois, "IBM, Amazon, Google and Microsoft partner with White House to provide compute resources for COVID-19 research," T*ech Crunch,* March 22, 2020, https://techcrunch-com.cdn.ampproject.org/c/s/techcrunch.com/2020/03/22/ibm-amazon-google-and-microsoft-partner-with-white-house-to-provide-compute-resources-for-covid-19-research/amp/.

[28] Ashley Johnson, "How Artificial Intelligence is Aiding the Fight Against Coronavirus," *Center for Data Innovation*, March 13, 2020, https://www.datainnovation.org/2020/03/how-artificial-intelligence-is-aiding-the-fight-against-coronavirus/.

[29] Martin Pollard, "Even mask-wearers can be ID'd, China facial recognition firm says," *Reuters*, March 9, 2020, https://www.reuters.com/article/us-health-coronavirus-facial-recognition/even-mask-wearers-can-be-idd-china-facial-recognition-firm-says-idUSKBN20W0WL.

[30] "The real cost of a business cyber attack. The cost of a cyber-attack vs cyber insurance," *Vista Insurance*, November 2021, http://www.vistainsurance.co.uk/10-largest-cyber-attacks-2020/.

[31] Juta Gurinaviciute, "5 biggest cybersecurity threats. How hackers utilize remote work and human error to steal corporate data," *Security Magazine*, February 3,

2021, https://www.securitymagazine.com/articles/94506-5-biggest-cybersecurity-threats.

[32] James Vincent, “Facebook is now using AI to sort content for quicker moderation. A little more machine learning in the moderation mix,” *The Verge*, Nov 13, 2020, https://www.theverge.com/2020/11/13/21562596/facebook-ai-moderation.

[33] “Industry’s First Cyber Security Platform with Autonomous Threat Prevention,” *Globe Newswire*, Nov. 09, 2020, https://www.globenewswire.com/news-release/2020/11/09/2122492/0/en/Check-Point-Software-Launches-Industry-s-First-Cyber-Security-Platform-with-Autonomous-Threat-Prevention.html.

[34] “Machine Learning: Higher Performance Analytics for Lower False Positives. Effectively Fight Financial Crime with Analytical Agents Trained on Large Data Sets,” *Verafin*, August 2, 2019, https://verafin.com/2019/08/machine-learning-higher-performance-analytics-for-lower-false-positives/.

[35] Tushaar Gangavarapu, C. D. Jaidhar, and Bhabesh Chanduka, “Applicability of machine learning in spam and phishing email filtering: review and approaches,” *Artificial Intelligence Review* 53 (2020): 5019–81, https://link.springer.com/article/10.1007/s10462-020-09814-9.

[36] Derek Lin, “A Machine Learning Study on Phishing URL Detection,” *Exabeam*, May 05, 2017, https://www.exabeam.com/information-security/machine-learning-study-phishing-url-detection/.

[37] Derek Lin, “Thorough Analysis For Using Data Science To Detect Malicious,” *Exabeam*, September 09, 2015, https://www.exabeam.com/information-security/thorough-analysis-for-using-data-science-to-detect-malicious-domains/.

[38] “Artificial Intelligence Cybersecurity Challenges,” ENISA, December 15, 2020, https://www.enisa.europa.eu/publications/artificial-intelligence-cybersecurity-challenges.

[39] “Ethics guidelines for trustworthy AI,” ENISA, 08 April 2019, https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai.

[40] “Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self-assessment,” European Commission, 17 July 2020, https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment.

[41] Tom Bur, “New Steps to Combat Disinformation,” Microsoft, Sep 1, 2020, https://blogs.microsoft.com/on-the-issues/2020/09/01/disinformation-deepfakes-newsguard-video-authenticator/.

[42] Ibid.

[43] Nicolas Papernot, Shuang Song, Ilya Mironov, Ananth Raghunathan, Kunal Talwar, and Úlfar Erlingsson, “Scalable Private Learning with PATE,” *arxiv*, 4 Feb 2018, https://arxiv.org/abs/1802.08908.

[44] Nathan Dowlin, Ran Gilad-Bachrach, Kim Laine, Kristin Lauter, Michael Naehrig, and John Wernsing, “CryptoNets: Applying Neural Networks to Encrypted Data with High Throughput and Accuracy,” *Proceedings of the 33 rd International Conference on*

Machine Learning, New York, NY, USA, 2016, http://proceedings.mlr.press/v48/gilad-bachrach16.pdf.

[45] Alex McFarland, "Early Warning System for Disinformation Developed with AI," *Unite.AI*, April 7, 2020, https://www.unite.ai/early-warning-system-for-disinformation-developed-with-ai/.

[46] "Scientists voice concerns, call for transparency and reproducibility in AI research," *Science Daily*, October 14, 2020, https://www.sciencedaily.com/releases/2020/10/201014114606.htm.

[47] "Security software for autonomous vehicles," *Science Daily*, September 16, 2020, https://www.sciencedaily.com/releases/2020/09/200916113601.htm.

[48] "How to make AI trustworthy. New tool might aid the adoption of technologies such as autonomous vehicles," *Science Daily*, August 31, 2020, www.sciencedaily.com/releases/2020/08/200827105937.htm.

[49] Jennifer Bresnick, "Deep Learning Malware Can Fake Cancer on Medical Images. A deep learning algorithm successfully penetrated a healthcare organization and fooled both humans and an AI system with faked medical images," *HealthITAnalytics,* April 05, 2019, https://healthitanalytics.com/news/deep-learning-malware-can-fake-cancer-on-medical-images.

[50] Rob Toews, „Deepfakes Are Going To Wreak Havoc On Society. We Are Not Prepared," *Forbes*, May 25, 2020, https://www.forbes.com/sites/robtoews/2020/05/25/deepfakes-are-going-to-wreak-havoc-on-society-we-are-not-prepared/.

[51] Zoe Schiffer "Surveillance company harassed female employees using its own facial recognition te,chnology. Verkada's clients include Juul Labs, Equinox, and Red Lobster," The Verge, Oct 26, 2020, https://www.theverge.com/2020/10/26/21535089/surveillance-company-verkada-harassed-female-employees.

[52] Bruce Schneier, "AI Has Made Video Surveillance Automated and Terrifying," *Belfer Center*, June 13, 2019, https://www.belfercenter.org/publication/ai-has-made-video-surveillance-automated-and-terrifying.

[53] Khari Johnson, "Clearview AI stops facial recognition sales in Canada amid privacy investigation," *Venture Beat*, July 6, 2020, https://venturebeat.com/2020/07/06/clearview-ai-stops-facial-recognition-sales-in-canada-amid-privacy-investigation/.

[54] Adam Hadhazy, "Protecting smart machines from smart attacks," Princeton University, Oct. 14, 2019, https://www.princeton.edu/news/2019/10/14/adversarial-machine-learning-artificial-intelligence-comes-new-types-attacks.

[55] Ajitesh Kumar, "Security Attacks Analysis of Machine Learning Models," *Data Analytics*, October 6, 2018, https://vitalflux.com/security-attacks-machine-learning-models/.

[56] World Bank, "Harnessing Artificial Intelligence for Development on the Post-COVID-19 Era: A Review of National AI Strategies and Policies," 2021, https://openknowledge.worldbank.org/handle/10986/35619.

[57] OECD (Organisation for Economic Cooperation and Development), "Using artificial intelligence to help combat COVID-19" report, 2020, https://www.oecd.org/corona

virus/policy-responses/using-artificial-intelligence-to-help-combat-covid-19-ae4c5c21/.

[58] Jeremy Kahn and Jonathan Vanian, "How A.I. can speed up the COVID-19 vaccination drive," *Fortune,* 5 January 2021, https://fortune.com/2021/01/05/a-i-covid-19-vaccination-drive/ (accessed 1 March 2021).

[59] World Bank, "Harnessing Artificial Intelligence."

About the Author

George Sharkov is CEO of the European Software Institute CEE since 2003 and Head of the Cybersecurity Lab at Sofia Tech Park. He was an adviser to the Bulgarian Minister of Defense (2014-2021) and National Cybersecurity Coordinator, leading the national Cyber Resilience Strategy development. Member of the EU AI High Level Expert Group, SMEs voice at ETSI Technical Committees CYBER and ISG "Securing AI," ENISA Group on Secure AI, ENISA Stakeholders Cybersecurity Certification Group. He holds PhD in AI and is lecturing at four leading universities (software quality, cybersecurity and resilience, and active security). https://orcid.org/0000-0001-5086-311X