CLCT

WILLIAM & MARY
LAW SCHOOL

Cybersecurity and Information Security Newsletter

Issue 25 | April 21, 2023

Table of Contents

Daniel Shin, research scientist with the Commonwealth Cyber Initiative (CCI) Coastal Virginia region, wants to hear from you! Submit any cybersecurity and information security news items or request related topics, via e-mail to dshin01@wm.edu.

This newsletter supports the mission of CCI. To learn more about CCI, including upcoming events, funded research, and news, please visit cyberinitiative.org.

The White House publishes the 2023 National Cybersecurity Strategy

On March 1, 2023, President Biden released the *2023 National Cybersecurity Strategy* (Strategy), which provides a high-level overview of the U.S. government's goals and approach to secure the cyberspace domain. *FACT SHEET: Biden-Harris Administration Announces National Cybersecurity Strategy*, available here. The Strategy seeks to shift the responsibility to defend cyberspace away from individuals, small businesses, and local governments to organizations and entities "that are most capable and best-positioned to reduce risks" for the nation. *2023 National Cybersecurity Strategy,* available here. It also calls for pursuing long-term investments to bolster infrastructure supporting cyberspace against threats emerging on the horizon.

The newly published Strategy replaces the *2018 National Cyber Strategy* but retains some of the previous administration's priorities in place.

Summary

The Strategy's main goals are to build a defensible, resilient, and value-aligned digital ecosystem. *Id.* at 5. Specifically, the U.S. government seeks to develop a digital ecosystem where:

(1) cyber defense is overwhelmingly easier and cheaper to implement while being more effective,

(2) cyber incidents and errors have limited scope and lasting impact, and

(3) the values embodied by the Declaration for the Future of the Internet and the Freedom Online Coalition shape and reinforce the digital world.

To achieve these goals, the Strategy provides five "pillars" of actionable objectives that seek collaboration among digital ecosystem stakeholders to advance U.S. interests.

The Five Pillars

The first pillar is "**Defend Critical Infrastructure**," which focuses on the U.S. government's commitment to the availability and resilience of the nation's critical infrastructure and other essential services serving the public. This objective entails the establishment of relevant cybersecurity standards regulating critical infrastructure operators. Critical infrastructure refers to infrastructure that provides essential services so vital to the United States that their incapacitation or destruction would have a debilitating effect on security, national economic security, or national public health or safety. *Critical Infrastructure Sectors [CISA]*, available here; *Presidential Policy Directive 21*, available here.

To minimize the cost and burden of compliance, this pillar aims to harmonize existing regulations, and assessments and audits of regulated entities. *Id.* at 13. It also calls for deepened strategic public-private collaboration to facilitate critical infrastructure security and resilience. Finally, the U.S. government commits to long-term efforts to defend and modernize federal systems, including implementing zero trust principles.

The second pillar is "**Disrupt and Dismantle Threat Actors**," which involves coordinated efforts by federal and non-federal entities to proactively disrupt malicious operations conducted by threat actors, including other foreign government actors, criminal groups, and other entities. To achieve this aim, the U.S. government's efforts will focus on enhanced collaboration with public and private sector partners to improve intelligence sharing, perform cyber disruption campaigns, prevent adversaries from utilizing U.S.-based infrastructure, and disrupt global ransomware campaigns. Under this pillar, select private sector partners are encouraged to participate continually in collaborative disruption operations with the U.S. government to thwart malicious cyber efforts by threat actors. Finally, the Strategy recommends strongly that victims of ransomware attacks not pay ransom to threat actors but instead report the cyber attacks to the federal government for potential victim support.

The third pillar is "**Shape Market Forces to Drive Security and Resilience**," which focuses on placing the responsibility for reducing cyber incident risk "on the stakeholders most capable of taking action to prevent bad outcomes, not on the end-users that often bear the consequences" arising out of cybersecurity incidents. *2023 National Cybersecurity Strategy, supra* at 25. The pillar calls for enhanced data protection legislation to protect personal data and new legislation—in consultation with both Congress and the private sector—to establish liability for insecure software products and services. *Id.* 24-25.

The fourth pillar is "**Invest in a Resilient Future**," which continues the federal government's commitment to invest in secure and resilient next-generation technologies and infrastructure. Among many possibilities, the pillar commits to the prioritization of investments in post-quantum resilient transitions, where IT systems would be prepared for against threats introduced by quantum computing. (Due to quantum computing's next-generation computational power, current security mechanisms, including certain encryptions, can easily be defeated by quantum computers. *See Post-Quantum Cryptography Initiative [CISA]*, available *here*.)

Also, the Strategy calls for investments in digital identity solutions (e.g., mobile drivers' licenses) that provide robust security while maintaining privacy, civil liberties, equity, accessibility, and interoperability features.

The fifth pillar is "**Forge International Partnerships to Pursue Shared Goals**," pursuant to which the U.S. government will use international collaboration with its nation's partners to build a more resilient global digital ecosystem to guard against cyber threats. That would include leading a North Atlantic Treaty Organization's (NATO) effort to create a NATO cyber incident support capability that can assist NATO member states in guarding against significant malicious cyber operations. The pillar also includes a commitment for the U.S. to work with international partners to create norms of responsible state behavior that would govern malicious activities below the threshold of armed conflict. Finally, the Strategy calls for securing the global supply chain for information and communications technology and operational technology products and services.

Some of the National Cybersecurity Strategy components have already been implemented through the 2022 National Security Strategy, Executive Order 14028 (86 FR 26633), National Security Memorandums (National Security Memorandum 5 & National Security Memorandum 10), and Office of Management and Budget Memorandum M-22-09. The White House designated the Office of the National Cyber Director to implement the rest of the Strategy across the U.S. government.

Analysis

The Strategy's five pillars cover defensive and offensive cyber postures, potential changes to the market and legal cyber landscape, investments in newer technologies and critical infrastructure and global efforts for a more resilient digital ecosystem. Overall, it contains comprehensive approaches to meet current challenges while making commitments to address long-term issues on the horizon. For example, the fourth pillar, "Invest in a Resilient Future," addresses the current issue of digital identity security—which may include not only mobile drivers' licenses but also a potential replacement for social security numbers and other sensitive identification information—and issues on the horizon, such as developing quantum-resistant encryption.

Although many of the strategic approaches aim to create a stable, secure, and reliable digital global ecosystem, "Disrupt and Dismantle Threat Actors" contemplates certain offensive cyber strategies that may create more disruptions to the global digital infrastructure. Under this pillar, the U.S. government is committed to coordinated, proactive cyber operations to disrupt malicious acts initiated by threat actors, which may include *foreign government* actors. This coordinated offensive includes the strategic objective of enhancing public-private operational collaboration, which the Strategy envisions the U.S. government to develop a model for trusted private sector partners to engage offensively against threat actors.

The Strategy notes the 2021 takedown of the Emotet botnet, which involved the efforts of the U.S. government, other U.S. allied countries, and private industry to disrupt the botnet's operation. *2023 National Cybersecurity Strategy*, *supra* at 19. Citing the "interest of the cybersecurity community and digital infrastructure owners and operators in continuing this approach," the U.S. government calls for a sustained and expanded public and private collaboration model for disrupting malicious operations on a continuing basis. *Id.*

Despite the success of the Emotet botnet takedown, involving the private sector in conducting offensive cyber operations may implicate legal issues, especially if the targeted threat actor is a foreign state. Although the applicability of laws of war within the cyber domain is still under development, if a private company engages in offensive cyber operations against a malicious foreign state actor, the private company's conduct could be interpreted as an armed attack on behalf of the government.

Under the International Court of Justice's judgment in *Nicaragua v. United States*, an armed attack not only encompasses acts by regular armed forces but also "the sending by or on behalf of a State of armed bands, groups, irregulars or mercenaries, which carry out acts of armed force against another State." *Military and Paramilitary Activities in and against Nicaragua (Nicaragua v. United States of America), Judgment of 27 June 1986 [International Court of Justice]* at 93, available here.

There are fundamental questions about whether an offensive cyber operation could be recognized as an "armed attack" because cyber operations rarely cause kinetic (physical) damage. Even if cyber operations result in physical destruction, there is no international consensus on the applicability of laws of war in the cyber domain. Despite these legal uncertainties, directly involving private actors to engage in offensive cyber operations (especially against a state threat actor) may encourage other nations, including those who are hostile to the U.S., to hire private hacking groups to engage in cyber attacks against the U.S. or U.S. entities. This could threaten the global digital ecosystem.

Alibaba's DAMO Vision Intelligence Lab releases publicly Text-to-Video Diffusion Model

In March 2023, DAMO Vision Intelligence Lab of Alibaba (a China-based technology e-commerce company) published publicly a text-to-video diffusion model, allowing users to synthesize low-quality and short videos based on a text prompt using a high-end consumer computer or cloud computing environment. *damo-vilab/modelscope-damo-text-to-video-synthesis*, available here. Researchers published the diffusion model with accompanying code on two publicly accessible code repository platforms, allowing other users to download and test the model. *Text-to-video-synthesis Model in Open Domain [Hugging Face]*, available here; *damo/text-to-video-synthesis [ModelScope]*, available here.

To develop the text-to-video diffusion model, Alibaba researchers used both image and text pair data sets (LAION5b) and ImageNet) and a video and text pair data set (WebVid-10M) to train the model. Because the training data set relies on the English text description of the accompanying media, this model is able to synthesize video based on an English text prompt.

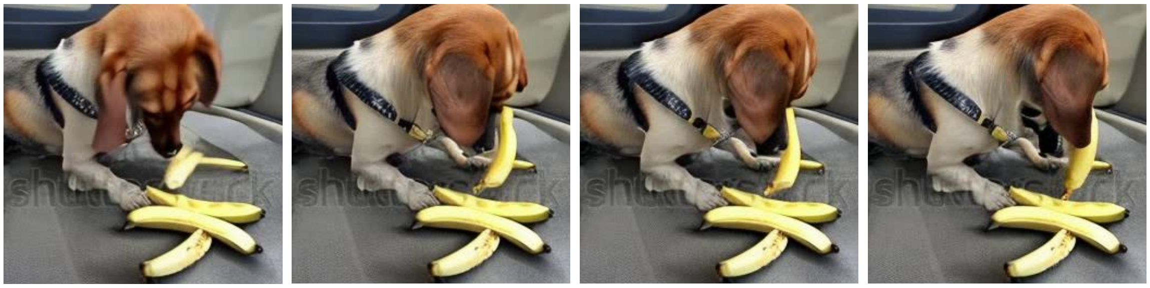

Figure 1: Sequential video frames of a synthesized video clip using the text-to-video diffusion model. The text prompt for the video was "A dog eating banana in an airplane." [Video available here]

To make it easy for other researchers and AI enthusiasts to evaluate the model, the researchers provided detailed instructions on how to use the published model and even provided an easy-to-use prepared code (i.e., Jupyter notebook) hosted on Google's cloud computing environment to allow non-technical users to experiment with the model online.

The text-to-video diffusion model is able to synthesize short, low-quality video clips based on a user prompt. The user prompt is the textual description of a video clip that the model attempts to synthesize. Given the complexity and the number of video frames necessary to generate a short video, even the most powerful consumer-level computers and cloud environments struggle to prepare a four-second video clip. Although the model is partially successful in attempting to synthesize video clips based on the user's prompt, the model sometimes ignores part of the user's description. Finally, generated video clips sometimes depict bizarre renditions where certain features within the video scene are exaggerated or unrealistic.

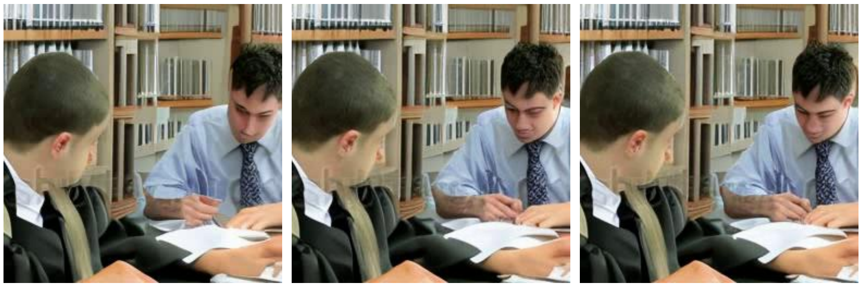

Figure 2: Video frames of a synthesized video clip using the prompt "A man and woman studying in a law school." Note that the model failed to follow the prompt fully by producing a video scene with two men. [Video available here]

Despite some limitations, Alibaba's researchers made profound progress in video synthesis within the diffusion model development. Unlike other publicly available diffusion-based video creation techniques, the text-to-video model is able to create video scenes based on the user-provided text description of the entire video, as opposed to a description based on individual video frames. This model is able to produce temporally consistent video clips without video flickering and scene variations that are usually present in other diffusion-based video clips. Because the model is trained on video clips and not on individual images (i.e., video frames), its synthesized output is properly tailored towards a sequence of images that coherently presents a moving picture instead of disjointed sequences of images.

One of the major sources of video clips used by WebVid-10M is sample stock video clips by Shutterstock, a company specializing in providing licensed stock photos and videos. Because Shutterstock's sample video clips contain a text watermark (the text "shutterstock" is transparently placed over the sample video clips), many of the synthesized video clips also retain Shutterstock's watermark. Although it is possible to synthesize videos without the watermark, it would likely take multiple tries to get an unwatermarked video clip that also reasonably depicts a scene described by the user-provided prompt.

Researchers have restricted the model use for non-commercial purposes, and they have specifically prohibited the model from being used for generating (1) content that is demeaning or harmful to people or their environment, culture, religion," (2) "pornographic, violent and bloody content," and (3) false or erroneous information. The model description page also lists the model's limitations and biases, both on the text prompt input and the synthesized video output. *damo/text-to-video-synthesis [ModelScope]*, *supra*.

Analysis

Since Stability AI's public release of its Stable Diffusion model, AI enthusiasts have been able to synthesize all types of images using their computers or cloud computing platforms. Although the availability of AI models allows for a democratization of cutting-edge technologies, the public release of AI assets forgoes model use restriction mechanisms that could have curtailed unethical and unlawful uses. For example, Stable Diffusion 1.5 model has been used to synthesize photorealistic nudity and deceptive images (in furtherance of a fraudulent act). Some AI enthusiasts have even "fine-tuned" successfully the original Stable Diffusion 1.5 model to synthesize even more detailed and complex imagery based on a targeted concept, individual, or place. There are even modified Stable Diffusion models that can synthesize images based on text prompts and human pose mapping inputs to tailor the characteristics of the final image.

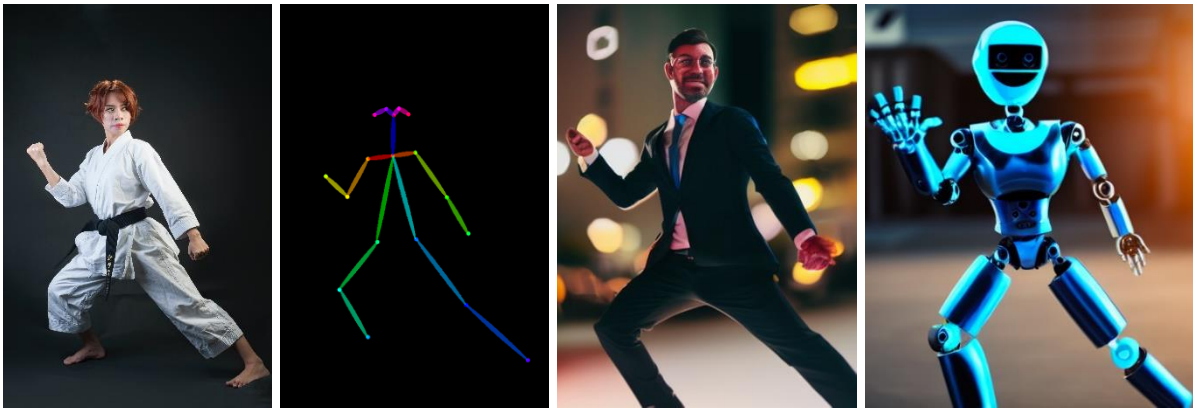

Figure 3: Stable Diffusion synthesis using the human pose information (second image to the left) from the original image (first image to the left). The synthesized images (first and second images to the right) incorporate both the user's text prompt and pose data of the original image.

Alibaba's text-to-video model is not able to synthesize the same level of realism and resolution as Stable Diffusion, but the model's public release allows the AI enthusiast community to explore novel video synthesis techniques and potentially develop newer features to the model.

Complementing the democratization of AI technologies, robust mechanisms should be in place to ensure that misuse of such technologies is minimized under ethical and legal norms. For generative AI models, researchers and developers may scrutinize the training datasets used to develop the model and check whether the training data contains unsafe information. Such practice would mitigate risks stemming from AI users generating unsafe content using the model. Researchers and developers may also consider third-party ethical and legal considerations before releasing publicly an AI model.

Ultimately, there needs to be an AI risk management framework tailored towards generative AI models that can inform developers of the risks of releasing AI technologies to the public. (Appendix B of NIST's AI Risk Management Framework notes that "challenging risks related to generative AI" is not comprehensively addressed by the publication. *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, available here.)

**This newsletter issue would especially like to thank Dr. Yan Lu of Old Dominion University for verifying information on DAMO Vision Intelligence Lab's publication.*