Written Testimony
Hearing on "Modern Scams: How Scammers Are Using Artificial Intelligence & How We Can Fight Back"
Senate Special Committee on Aging
Tom Romanoff
Director, Technology Project at the Bipartisan Policy Center
November 16, 2023

Chairman Casey, Ranking Member Braun, and all members of the Special Committee on Aging,

Thank you for inviting me to testify on this important and evolving topic of AI in Frauds and Scams.

My name is Tom Romanoff, and I am the Director of the Technology Project at the Bipartisan Policy Center. As director, I lead the organization's research and advocacy for bipartisan solutions in the technology sector. Our portfolio includes content moderation, data privacy, digital divide issues, and Artificial Intelligence. The latter is where we kicked off our efforts in technology when, in 2018, we partnered with Representatives William Hurd and Robin Kelly to create a National Strategy for Artificial Intelligence. We helped pass this strategy as House Resolution 1250 alongside seven other bipartisan sponsors. Outside of the Representatives who co-sponsored this work, we also engaged hundreds of experts across civil society, the private sector, and academia.

Prior to my role at the Bipartisan Policy Center, I advised Chief Information Officers and Chief Information Security Officers across the federal government on emerging technologies and compliance with existing regulations. Included in my clients were the Office of Management and Budget, the U.S. Food and Drug Administration, the General Services Administration, and others. In addition to acting as a senior advisor for these officials, I led teams that worked on critical federal-wide initiatives in technology modernization and cybersecurity, including the President's Management Agenda and the Federal Data Strategy.

I commend the leadership of this Special Committee for holding this hearing today, as this is a critical moment for policymakers to consider solutions that will better prepare us for the emerging role of AI in cybercrimes. I want to first level-set on how we got here. As a technology, AI has been around for a while. While its theoretical roots originated in the 50s, since then, data scientists and engineers have discovered increasingly sophisticated ways to leverage the technology to predict and identify data trends.

There are seven main branches of AI, ranging from robotics to natural language processing to deep-learning machines. All these branches of AI are still in the "Artificial Narrow Intelligence" development phase, meaning they can do a given task and are not as smart as a human across different functional areas. However, AI technology is increasingly being leveraged to support daily use that most users may not even recognize – from support with driving your vehicle to unlocking your smartphone.

Over the years, we have seen waves of interest in regulating this technology as new advances demonstrate new capacities. Generative AI, or the ability of a computer to produce content on its own, is the latest in this trend. Without a doubt, Generative AI has a significantly more demonstratable capacity than past AI breakthroughs because of its open (and free) accessibility and availability to most users.

Fast forward to today, Generative AI's capacity has gotten so good that most people cannot tell the difference between computer-generated content and human-generated content. This is due to a couple of factors, the most important of which is a *transformer.* In 2017, transformers were created to allow AI programs to get more done at a faster pace. Since that breakthrough, we have seen advanced capacity across all forms of AI. Robotics, computer vision, deep learning, and Generative AI have all gotten better at their tasks. In the Generative AI space, that means that we have seen advancements in creating, detecting, and accessing synthetic media, commonly referred to as deepfakes (images) or voice cloning. In other areas, it means that AI is also advancing; we do not have access to the output of those advancements freely available.

Generative AI is not inherently bad for our society or precluded from use in scams. In fact, many will argue that Generative AI has many more positive use cases than negative ones. In nearly every aspect of our lives, we can think of a way to deploy Generative AI to increase productivity and improve outcomes. Among aging Americans, for example, it can be used to detect elder abuse, address senior loneliness, and revolutionize medical care. The benefits of this technology have yet to be fully realized, but many in the tech space are working on products to leverage it for positive outcomes.

Despite these and many other benefits, we also know that Generative AI is already being used in cybercrime. Criminals are exploiting this technology to produce manufactured and realistic media. As the good in this technology is explored, we must acknowledge AI's risks and seek a balanced approach, focusing on curtailing abuse while promoting positive uses and innovation.

As a result of its current capacity, Generative AI increases the quality, quantity, and targeting capabilities of fraud-- making it cheaper, faster, and more effective to create idiomatic narratives, deploy multimedia resources, and write malicious code. Combined with opaque legal frameworks and international origins, criminals can now use GAI to coordinate sophisticated attacks with little risk of being caught. Examples include:

- Cybercriminals are leveraging generative AI to augment social engineering campaigns—cybersecurity attacks that use psychology to manipulate people into sharing sensitive information.
- Hyper-realistic voice deepfakes are used to manipulate victims and
- Fake nudes are used to extort in what is called "sextortion."

While many of the victims tend to be younger people, older people are targeted for more in absolute sums. Criminals know they have more money to lose.

Beyond the psychological trauma that comes with these kinds of crimes, the financial losses are skyrocketing. In 2020, Americans lost $3.5 billion to online fraud; by 2022, losses had tripled to nearly $9 billion (Federal Trade Commission). It is important to note that not all these scams have been perpetuated by AI technologies; conflating the total online fraud loss with the rise of GAI is a misrepresentative of the issue. The warning here is that GAI will make it easier and more prevalent as criminals adopt its use in their operations.

Importantly, we are seeing some indications that adoption is coming fast. A 2023 report released by the cybersecurity firm McAfee reported that GAI applications need less than three seconds of a person's recorded voice to produce a convincing clone. In a survey of more than 7,000 individuals worldwide, one in four said that they experienced an AI voice cloning scam. Even more telling is that the survey happened *before* ChatGPT was widely used, showing that voice-cloning technology has been in active use even before popular tools like ChatGPT 3 emerged.

For the elderly community, there are additional obstacles to navigating this new threat. In a survey this year, 68% of Gen X and Baby Boomers said that they do not use AI, with 88% in that demographic unclear about its impact on their lives. That is a problem if potential victims do not know of the technology or its applications.

We know that cyber fraud is a multi-billion-dollar-a-year business. As we navigate the evolving landscape of generative AI, it is imperative to recognize and address the emergent fraud risks associated with generating synthetic data at scale. Here are five risks that pose the greatest danger to older adults:

- Firstly, the creation of deepfakes and the dissemination of misinformation pose significant threats. Generative AI can fabricate highly realistic images, videos, and audio recordings, which can be used to mislead the public, manipulate opinions, and impersonate individuals for malicious purposes.
- Secondly, the rise in sophisticated phishing and social engineering tactics is alarming. AI-enhanced methods can mimic personal communication styles, making fraudulent emails and messages increasingly challenging to distinguish from legitimate correspondence, elevating the risk of individual and organizational data breaches.
- Thirdly, while identity theft has long been a concern, AI makes it much easier and more prevalent. Generative AI can produce authentic-looking images and documents, facilitating the creation of fake identities. This capability can be exploited in financial fraud, the creation of deceptive online personas, and circumventing security measures based on identity verification.

- Fourthly, manipulating financial markets through AI-generated misinformation is a looming threat. Fraudulent actors can use AI to fabricate news or social media content, influencing investor decisions and market trends for personal gain.
- Lastly, the automation of traditional scams, such as romance or lottery scams, has become more efficient and widespread due to generative AI. This amplifies the scale and reach of fraudulent activities, impacting a more significant number of victims and complicating the efforts to combat such schemes.

Solutions

In trying to prevent the use of GAI for crime, many are looking into using the technology to detect and mitigate scams before they result in financial loss. Suppose an AI can detect a Deepfake and label it; that could eliminate a lot of risk for the consumers of that media. It is promising but has some issues- criminals always look for ways to beat the detectors. While publicly available GAI may put a watermark in the digital background, the programs specifically developed to facilitate fraud will not, and they are getting more sophisticated. These could lead to an arms race between deepfake creators and detectors. Our current detectors are unreliable, especially if an adversarial AI is trained to beat it.

This is an area where we have seen a lot of interest from policymakers at both the state and federal levels. Many of the bills introduced in the House are intended to counter the deepfake risk. This is also one of the most popular state-legislature topics, with several states advancing bills to identify and govern synthetic media. Last week, President Biden signed an executive order with sweeping AI provisions, including directing the Commerce Department to:

1) Verify the originality and trace the origins of content.
2) Mark artificially created content, for instance, through watermark methods.
3) Identify artificially generated content.
4) Restrict AI systems from creating content that depicts child sexual abuse.
5) Evaluate tools employed for the tasks mentioned above.
6) Review and manage synthetic content.

Additionally, the Executive Order mandates that NIST provides instructions to federal agencies, particularly concerning the labeling and verification of content they create or disseminate. One of the outstanding questions that needs to be addressed is the definition of synthetic media: does any altered media fall into the category of synthetic media, and how do you distinguish between AI-generated media and replicated media? Another area of consideration is the role that Section 230 has in liability protections for platforms that post these images.

AI does have a role in countering these issues. The only way to process the sheer amount of information and identify patterns across the US criminal network is through a program that can accurately and fairly predict trends. AI will be at the center of that program. Companies who invest in these systems will reap the rewards of lower operational costs and reduced liability. AI can amplify the defense of these systems.

Another way we can address this risk is to adjust and adopt multiple authentication factors in the methods used to validate individuals' identities. For example, some banks now use voice identification to authenticate account ownership. However, criminals have used AI to clone voices to break into a bank's voice banking system. Using multiple authentication processes to validate access and secure assets can address this issue. Biometrics cannot be the only authenticator but can be part of a system of protection.

Further, those creating AI systems must address bias in the data used to train the system, particularly who gets flagged as a potential criminal. For instance, companies have been fined for discrimination in their lending or credit decisions based on zip codes, name suffixes, or other indicators. These are human biases. If those same applications were used to train an AI to predict fraud, the bias and discrimination would be built into a system and scaled up as an organization deploys this program.

In closing, I am leaving you with some reference to which this technology is being adopted and is advancing. In the ten years, this tech has been around, it has gone from computer science theory to widescale use. ChatGPT reportedly hit 100 million users in February after two months of free and public access, an unprecedented technology adoption. Over the last year, AI has been integrated into almost every major tech company's platform, and we have seen many new uses, including in cybercrime, emerge. In the coming months, we will see billions in start-up funding go toward building out this technology's use. Do not let the idea that this is tomorrow's technology cloud the reality of its use today. As I stated in the beginning, we must tackle the abuse while driving toward positive applications to safeguard its adoption.