ICANN82 | Foro de la comunidad – Debate de NARALO sobre la IA y el uso indebido del DNS
Lunes, 10 de marzo de 2025 – 15:00 a 16:00 PST

MICHELLE DESMYTER: Bienvenidos todos. Vamos a comenzar en unos minutos. Vamos a comenzar en breve. Por favor, nos vamos sentando. Va a comenzar la sesión. Comenzamos a grabar. Muchas gracias.

Hola y bienvenidos todos a la mesa redonda de NARALO: Uso indebido del DNS e IA. Combate y tendencias. Soy Michelle DeSmyter y soy la coordinadora de participación remota de la reunión. Tengan en cuenta que la reunión se está grabando y se rige por los estándares de comportamiento esperado de la ICANN y la política de antiabuso de la comunidad de la ICANN.

Las preguntas y comentarios se van a leer en voz alta solamente si se incluyen en la ventana de preguntas y respuestas. Habrá interpretación en inglés, francés y español. Si quieren tomar la palabra durante la sesión, por favor, levanten la mano en Zoom. Cuando se diga su nombre los participantes virtuales van a recibir permiso para habilitar el micrófono y los que estén presentes tendrán que utilizar el micrófono físico. Digan su nombre para los registros, el idioma en el que van a hablar si no es inglés y, por favor, hablen a una velocidad razonable. Le paso la palabra a Greg Shatan, el presidente de NARALO.

GREG SHATAN: Gracias por venir a nuestra mesa redonda sobre uso indebido del DNS e inteligencia artificial. Varios temas muy adecuados para estos tiempos. Sin más, dado que tenemos un panel de luminarias voy a presentarles primero directamente a mi derecha, Jeff Bedser. Es un empresario líder en DNS, en mitigación. Es el CEO de CleanDNS, una empresa que tiene la misión de limpiar Internet para siempre deteniendo la actividad maliciosa en el DNS. En su última vida fue detective privado y utiliza su conocimiento para proteger a los usuarios finales. La gente que nos importa en At-Large. Está en distintos comités sobre gobernanza, sobre jurisdicción y tiene 25 años de experiencia en el campo. Fundó iThreat, una empresa contra las amenazas en Internet. Trabaja en SSAC en ICANN desde 2007 y formalmente estuvo en la junta del PIR. Gracias por estar con nosotros, Jeff.

Tim Maurer a su derecha. Es el director sénior de ciberseguridad en Microsoft, donde colabora estrechamente con los equipos y partes interesadas, y también las externas en Microsoft. Trabajó para el gobierno previamente como consejero sénior de ciberseguridad en las tecnologías de seguridad de interior y luego en el Consejo de Seguridad Nacional de la Casa Blanca, en la gestión del Presidente Biden. Viene trabajando en política de ciberseguridad y tecnología. Es el autor de *Cyber Mercenaries: The State, Hackers, and Power*.

A su derecha tenemos a Ana Neves. Muchos de ustedes probablemente la conocen como representante del GAC de

Portugal. Ella tiene distintos trabajos. Es miembro del IGF MAG actual, que es una junta que mantiene funcionando el DNS. Dirige la junta asesora en FCT, FCCN MECI en Portugal. También es vicepresidenta de la Comisión de la ONU de Ciencia, Tecnología y Desarrollo. Tiene 30 años de experiencia en política pública, desarrollo tecnológico. En 2018 recibió el premio como líder europea en su sector. Tenemos un panel que me hace a mí parecer oscuro y pequeño.

Sin más, quisiera comenzar el debate. Si podemos comenzar con la primera diapositiva. Antes de avanzar con esto quisiera mencionar que toda la presentación fue escrita por Microsoft Copilot. No es broma. Ahí está. En lugar de que el Señor sea mi copiloto, Microsoft es mi copiloto.

Según Copilot, el uso indebido del DNS hace referencia al uso indebido del DNS con objetivos maliciosos tales como phishing, malware y otras amenazas. La definición oficial de la ICANN del uso indebido del DNS tiene cinco categorías, una por cada región geográfica. Malware, botnets, phishing, pharming y spam, pero solamente cuando es un mecanismo de entrega de las otras formas. El spam per se no es uso indebido del DNS.

Aquellos que ya hace tiempo que estamos aquí venimos hablando sobre los límites y las definiciones del uso indebido del DNS y en lugar de seguir hablando del mismo vamos a vivir con esta definición del uso indebido del DNS a medida que avanzamos y ya nos da problemas suficientes.

Como saben ustedes, nosotros y también Copilot, la IA y el uso de enseñanza por máquinas se está utilizando para combatir esta amenaza detectando y mitigando distintas amenazas de manera más efectiva. Eso léanlo ustedes cuando tengan tiempo. No hace falta escucharlo aquí de los panelistas o de lo que tiene que decir Copilot.

El diseño fue también elegido por mí, por el PowerPoint de Microsoft. La imagen de la portada también la eligió Copilot. Soy muy creativo pero decidí que Copilot hiciera todo. En cualquier caso, quisiera en primer lugar pasarle la palabra a Tim.

Un minutito más. El uso indebido del DNS también se ve manifestado en la inteligencia artificial y hay varios ejemplos en esta presentación de la manera en que la IA está creando nuevas formas y mutando las formas más antiguas del uso indebido, por lo cual tenemos una especie de batalla de la IA hacia el bien y hacia el mal.

Nuestros tres panelistas van a estar comentando en general la interfaz IA y la ciberseguridad, cómo afecta al DNS y los efectos en general. Esto es una reseña general y el único que está internamente en este panel es el que va a empezar a tomar la palabra.

TIM MAURER: Es un placer estar aquí. Gracias, Greg. He estado hablando con Ana y creo que la última vez que nos vimos fue antes de la pandemia, la

última vez que pude venir a una reunión de la ICANN. Es un placer estar con todos ustedes hoy porque es una comunidad que nosotros esperamos que nos ayude a maximizar la defensa del DNS y lograr lo que esperamos que libere la tecnología, que es el progreso ulterior humano, y distintas maneras que nos ayuden a avanzar en metas y objetivos que compartimos que existen hoy.

Hace más de un año que estoy con Microsoft y ha sido una experiencia maravillosa para mí estar en esta empresa y al frente de la innovación en IA. La manera en la que estamos pensando más abarcatoriamente, pero también específicamente el uso indebido del DNS, es que sabemos que la IA hace un tiempo que está, según como la definamos. La predictividad y el aprendizaje por máquina son maneras en las que se usó y también tenemos que pensar en el uso indebido del DNS. IA generativa nos va a ayudar de distintas maneras y también ayudarnos a mitigar de manera más accesible el uso indebido del DNS, incrementando los niveles de habilidades y con Copilot y otras herramientas, para que se puedan utilizar herramientas de manera más efectiva en contra del uso indebido del DNS.

Me gustaría ser lo más interactivo posible. La empresa considera IA en el contexto de seguridad y uso indebido del DNS a través de tres lentes. Primero, cómo maximizar IA en defensa. Cómo movilizamos a los clientes y a la comunidad, tal como es la comunidad de la ICANN, para adoptar la tecnología con rapidez.

Hablo de velocidad porque les puedo garantizar que aquellos que tienen propósitos oscuros van a ser los que más rápido van a ver cómo pueden utilizar las herramientas para detectar la actividad maliciosa. Cuanto más rápido podamos avanzar la IA y la podamos utilizar para defensa y mitigación, más rápido podremos no solamente estar a la altura o quizá adelantarnos a las amenazas que están golpeando a las organizaciones de todo el mundo todos los días.

El segundo tema es que la empresa está muy focalizada en ver los pasos que tenemos que dar para que no se puedan aprovechar de nuestros productos y servicios. Hace un año, la empresa emitió una política que aclaraba que si detectábamos una amenaza, y hay muchos a los que estamos haciendo seguimiento en estos años. Hay quien intenta un acceso no autorizado o que nos llame la atención, que veamos que utiliza IA, eliminaremos su cuenta. Ya sea que la utilice de manera productiva.

No queremos que un delincuente sea más productivo. Queremos tomar acción. Esa es una política de los líderes no en respuesta a una acción gubernamental sino un compromiso para verificar que empoderemos a quien defiende el bien y le dificultemos la actividad a los delincuentes.

El último tema, como un marco general, es que la empresa está viendo lo que hace faltar hacer de nuestro lado para verificar que la tecnología se utilice en pos de nuestros servicios y también proteger productos y servicios que, por supuesto, se imaginarán

que tienen amenazas diversas, que conocen cómo los utilizamos. Sería un marco general, tridimensional. Sé que hay mucho que se puede hablar cuando hablamos del uso indebido del DNS en todo su detalle.

GREG SHATAN: Gracias, Tim. Tomé nota de que mencionó la defensa del DNS contra el uso indebido del DNS varias veces. Es un placer y tenemos la suerte de tener a Jeff Bedser aquí, cuya vida está definida por la defensa y limpiar el DNS. No necesitamos mostrar credenciales aquí. Jeff, en todo caso me gustaría que nos contara un poco. Sé que tiene diapositivas pero no fue escrito por Copilot. Sin más, le voy a pasar la palabra a Jeff.

JEFF BEDSER: No puedo negar que he utilizado el IA en muchas diapositivas pero no en esta. Voy a capitalizar el hecho de que hablé ya con este grupo hace una semana, creo, pero no, fue ayer, sobre el uso indebido del DNS. Creo que esta presentación elaborada tiene que ver con lo limpio del DNS y el uso de IA en procesos para mejorar esto como decía Tim.

Como mencioné, la presentación de ayer, si pensamos hoy, ellos ya lo pensaron ayer y más rápido, porque están delante de nosotros y no vamos a utilizar las tecnologías emergentes a la misma velocidad en defensa. Vamos a quedar fuera con tal velocidad.

Lo que voy a comentarles hoy es cómo venimos implementando de manera activa modelos de IP en lo que hacemos porque esto tiene que ver, como les decía el otro día, con que hay un volumen significativo de cosas que no conocemos y por qué no se están denunciando.

Hay denuncias de cinco a siete fuentes. La gente que denuncia solo representa la punta del iceberg. Hay muchas cosas que no se denuncian. La mayoría de la gente no habla de phishing, de smishing. Nadie denuncia. No nos enteramos. Entonces no podemos actuar al respecto.

Es decir, la mejor solución es la detección temprana, mitigación temprana. Estamos nosotros entre las empresas que detectan el uso indebido y la industria, nuestra industria, entre los registros, registradores y hosting para mitigarlo, para acelerar la denuncia, la detección para llegar al punto en el que el proceso entre la detección y la mitigación sea segundos.

Todos habrán leído que yo fui presidente de SSAC e interoperabilidad y uso indebido del DNS. Encontramos este estudio hace varios años. Había 96 horas entre la denuncia y la mitigación. Cuánta gente entra en 96 horas en un enlace maligno. Si podemos reducirlo a segundos en lugar de 96 horas, imagínense cómo mejoraría esto.

SAC115 fue la plantilla que se usó para el DNS. Muchas personas saben que aunque fue un papel muy bueno académico, a no ser que alguien construyera las soluciones en la cual indicaban en el

papel, no iba a ser útil. CleanDNS logró esto. Byron Holland utilizó esa terminología cuando hablamos de ello. No es solo resolver el problema sino que quienes lo reportan hagan un mejor trabajo y que la industria también mejore para solucionar el problema en el medio. Se encontró y qué hay que hacer para actuar sobre ello y después eliminarlo. Es un modelo agregado para escalar el proceso. Siguiente diapositiva, por favor.

Lo primero es fácil. Es un lugar donde uno puede rápidamente ver que la IA es mejor que lo humano. Todos me imagino que han recibido phishing. Si no es así, díganme por qué están aquí, en esta conferencia de tecnologías sin un celular. Nos pasa a todos.

El phishing ocurre porque hay un tipo de atracción. Atrae a las personas. Las personas piensan que están yendo a su banco local o a alguna institución que conoce. Usando el ejemplo del banco, uno piensa que entra al banco a nivel de Internet pero para los humanos, con el phishing el humano no conoce todos los bancos o todas las empresas del mundo. Los sistemas de IA pueden sacar esa información de la base de datos y tienen esas imágenes para analizar y comparar. Pueden validar el logo. En análisis de imagen pueden mirar la imagen y ver si hay un intento de capturar credenciales o información personal. Cuando yo hablé de análisis de imágenes como la captura de una pantalla, por ejemplo, una imagen de alguna finca, las siglas que se utilizan en las diferentes plataformas.

La IA es muy buena en ver el contenido y decir: “Por la estructura de esto, lo más probable es que sea phishing”. Al mirar el texto y otros componentes, disculpe que en la diapositiva no se ve muy bien lo que puse para describir estos esfuerzos. Disculpe.

Hablando de análisis de sentimiento, como lo llaman, es un análisis de informes de correos electrónicos que entran y comparan el contenido. Análisis de sentimiento, como lo llaman, toma los datos no estructurados de los correos electrónicos y ver qué es lo que está pasando. A los humanos les gusta contar un relato y les gusta contar por qué recibieron el phishing. Cuentan las historias de cómo fueron tramados.

Pueden ser historias fascinantes pero toma mucho tiempo. Lo que uno simplemente busca en el texto es cuál es la URL, cuál es el dominio y de qué es acusado. Recibió phishing y es una URL. Eso es lo único que importa y el resto de la historia no tiene nada que ver con los esfuerzos de mitigación. Las relaciones contextuales, disculpen que voy a dejar mis anteojos aquí, las incrustaciones de vectores, las similitudes entre los bloques de texto. Buscamos patrones en spam y phishing. Preguntamos si es la misma campaña del ciberdelincuente y si estamos utilizando la misma estructura para entender estos grupos más grandes.

Utilizar estas relaciones contextuales para averiguar que hay 10.000 de esto, vamos a removerlo de una vez porque todo está relacionado con lo mismo, especialmente cuando uno ve la hora o

el tiempo, y ve los datos y los compara. Siguiente diapositiva, por favor.

Me adelanté pero parte del análisis de imágenes, detección de logo e identidad, también detección de PII. Muchos dominios que se utilizan para el uso indebido lo ponen en una página de estacionamiento, una página de principio. Piensan que es una página simplemente de estacionamiento, lo que llaman parking page. Piensan que la página se irá pero se convierte en una página phishing. Cuando uno detecta si es una página estacional o si es una página de phishing, es la detección de la que estamos hablando. El punto es reducir el tiempo para poder analizar eso.

En el futuro queremos afinar este proceso. Es bueno para phishing estos esfuerzos pero qué otros daños podrían existir porque el uso del DNS, según lo definen las partes contratadas, es lo que digo Greg, pero hay otros tipos de daños que existen y se deben mitigar. Las partes que representan a las partes contratadas, como ccNSO y demás, tienen esos daños pero de pronto no pueden actuar en cuanto a ello. Queremos algo que cuando se vea un daño y actúe de esta manera, se defina como un daño y así mitigarlo.

Detección de código malicioso. Eso es difícil porque hay que procesar el código para determinar si es malicioso o no, entonces sus sistemas están en riesgo. Utilizar la IA y mirar los patrones para decir a qué familia de malware pertenece esto o cuál es el objetivo. Es una forma que se puede utilizar para combatir el malware.

Por último, el componente de la IA generativa. Esto lo hemos utilizado para crear plantillas y modelos de comunicación en base al tipo de uso indebido. Creamos las plantillas y hay que tener cuidado con esto porque la IA sin restricción empieza de pronto a crear otras entidades.

Por ejemplo, yo hice una investigación en algunas entidades y no existía ninguna de las llaves ni las referencias que hicieron. Uno tiene que tener cuidado con la IA generativa, para controlar el entorno y las restricciones que uno saca de allí. Siguiente diapositiva.

Fallas. Transparencia. Por ejemplo, no queremos hacer como si no estuviéramos utilizando estas tecnologías. También existen prejuicios de datos. Es una preocupación cuando tiene que ver con qué conjuntos de datos aplican. Por ejemplo, si hay algunos que tienen falsos positivos y no se sabe que son falsos positivos, entonces yo puedo desviar mis datos en cuanto a validación.

A veces estos conjuntos de datos si se confirman de una manera, entonces la parcialidad puede ser problemática. Estos falsos positivos pueden ser un problema. Tim va a hablar más adelante en cuanto a ese aspecto. Existe también el manejo de casos de una manera sensible y también manejarlo tomando en cuenta la privacidad de los datos. Creo que esa es la última diapositiva. Le doy la palabra de nuevo a Greg.

GREG SHATAN: Muchas gracias, Jeff. Muy interesante y a tiempo. Ahora le voy a dar la palabra a Ana Neves. Ana, puede decir lo que le dé la gana.

ANA NEVES: Muchas gracias. Si puedo decir lo que me dé la gana voy a decir que estoy feliz de estar aquí, en Seattle, especialmente aquí, en At-Large. Siempre es un gran honor el estar aquí con At-Large porque desde el principio de mi vida en el GAC, siempre defendí al GAC y At-Large, y que trabajaran en conjunto, porque debemos aumentar nuestros esfuerzos en el sector público y por el lado de los usuarios finales.

Qué puedo decir después de que hablaron Microsoft y Jeff. Explicaron lo que es el uso indebido del DNS. Lo que puedo decir de Microsoft es que es bueno saber que hay una política para mitigar el uso indebido del DNS pero qué pasa con otras empresas que no tienen esas políticas, y que esas políticas provienen de quien lo esté haciendo. ¿Podemos confiar en esas empresas para que tengas establecidas esas políticas? ¿Por qué?

Del sector público, que creo que por eso me invitaron a formar parte de esta mesa redonda, es para enfatizar la colaboración entre los sectores públicos y privados y tener más asociaciones y ver qué se puede hacer con los organismos reglamentarios y siempre tomar en cuenta la ciberseguridad y tener esto en cuenta cuando están regulando estas políticas públicas. Siempre tener el uso indebido del DNS en mente y ver que los esfuerzos de múltiples

partes interesadas también se incluyan porque es importante con respecto también al uso de IA en la seguridad del DNS.

También quiero resaltar otro punto que es el del desarrollo de los marcos de gobernanza de IA. Es bueno hablar de esto aquí, en ICANN y junto con At-Large, y quizá explorar algunas políticas donde la seguridad esté en conjunto con la gobernanza. También el repaso del CMSI+20 y también incluir el proceso del Pacto Digital Global y qué debemos ver en estos procesos y estas negociaciones.

Lo último que quiero resaltar es empoderar a los usuarios finales. Se debe diseñar IA con transparencia y permitir que los individuos puedan tomar decisiones informadas con respecto a la seguridad. Siempre se trata de tener buen sentido común pero que también los usuarios finales estén bien informados. Siempre es importante el desarrollo de capacidad, también la privacidad es importante en todos nuestros esfuerzos.

Yo diría que fortalecer la colaboración y tener mejores asociaciones y empoderar a los usuarios finales, esos serían los primeros temas para esta discusión y para esta mesa redonda. Muchas gracias.

GREG SHATAN: Gracias, Ana. ¿Podemos volver a las diapositivas donde habíamos interrumpido un momento? Vamos a la siguiente diapositiva para poder hablar de lo que nuestros amigos en Copilot encontraron con respecto a este tema. Hablar de otros hechos. Tenemos ataques de phishing impulsados por la IA donde el phishing se ve

más personalizado. Hacen la mímica como si fueran contactos confiables y traspasan los filtros de seguridad.

Creen que personas que son inteligentes no lo van a leer porque a veces sí existen errores de gramática y también la IA se puede utilizar para prejuzgar credenciales, hacerles la prueba a esas credenciales en segundos. Esas credenciales robadas. Hacen un ataque con fuerza bruta. La IA puede aumentar el poder del volumen para estas amenazas de actores malos.

El malware también puede ser aumentado por la IA con un código que se adapta dinámicamente para evadir la detección y ayuda a poder ocultarse de los antivirus porque va aprendiendo estos patrones y, claro, utilizando la IA. Como la guerra de los bots. Siguiente diapositiva, por favor.

Nosotros también podemos ver que se utilizan herramientas de escaneo de networking, identificar los puntos débiles en la infraestructura. Es como un tipo de ataques de penetración automatizados también. También ingeniería social. También puede convertirse en algo más dinámico, algo más real pero sin embargo falso, para entonces traspasar estos filtros de seguridad y utilizan la IA para hacer unos cambios para que no parezca como si fueran amenazas obvias. Esto claramente muestra por qué nosotros tenemos que utilizar características de defensas que continuamente estén mejorando. Siguiente diapositiva, por favor.

Algunas de estas soluciones también son herramientas impulsadas por la IA de seguridad con verificación estricta. Creo que muchos

de nosotros nos hemos dado autenticación de múltiples facetas, diferentes etapas, aun en los últimos seis meses al año, simplemente para estar al tanto de los niveles de amenaza.

Las plataformas de amenazas ampliadas, tal como comentó Jeff, siempre hay que considerar el factor humano. Hay que entrenar a gente real, empleados usuarios, público en general sobre cómo reconocer el phishing, los ataques de ingeniería social. Hay que concientizar a la gente. En mis empresas hacemos capacitaciones anuales. Siguiente.

Como abogado, soy consciente de la necesidad de incrementar los marcos constantemente, moverse no tan rápido como la IA pero tratando de mantener reglas legales y éticas de manera tal de hacer que la IA sea más responsable en cuanto a su uso.

Las empresas que están trabajando en IA, que sean conscientes del uso indebido del DNS y su universo en comparación con todos los otros universos que utilizan IA. Uno no siempre es consciente de estos espacios. Las empresas de IA están trabajando. No sé si lo suficiente como para poder compatibilizar con el resto del ecosistema.

Finalmente, trabajar con medidas más proactivas de ciberseguridad. También es importante que haya un grado de paranoia. Aquí tenemos la biografía de Jeff, que me dio Copilot de Microsoft. La pueden leer cuando quieran. Parece muy precisa y no inventó demasiado. Una diapositiva atrás, por favor.

Fíjense en el año en el segundo párrafo. 2071. Había una nota al pie que no eliminé. Ese error es humano. Siguiente, por favor. Este es el primer intento de IA. No mencionó a Microsoft para nada. No está actualizada pero sí un poquito más correcta. Siguiente diapositiva.

Agregué el prompt Microsoft y encontramos su cargo en la empresa. Si no hubiera agregado la palabra Microsoft en el prompt no lo hubiera encontrado. Copilot no conoce bien a Tim. Creo que ahora el problema se resolverá. Siguiente. Algo más de Tim. Sus hobbies. Andar en bicicleta, fotografía, lectura. Tiene distintos compromisos con el crecimiento personal y el bienestar. Eso es lo que cree Copilot. Siguiente.

Aquí tenemos la biografía de Ana. No parece completamente errónea. En el primer intento nos dio una Ana Neves totalmente distinta. Ana Luisa Neves. No utilicé su segundo nombre pero esta persona también estaba en nuestro mundo. Es directora de salud digital del Imperial College de Londres. Espero que en algún momento se conozcan. Búsquese. Quizá tengan cosas interesantes para hablar entre ustedes sobre el futuro digital. Siguiente.

Este soy yo. Si quieren saber algo de mí. Está un poquito antiguo. Ya he estado en más de 35 reuniones de la ICANN. Ahora soy secretario de la ISOC de Nueva York. Siguiente. Eso es todo. Esto es otra imagen de Seattle en la niebla, que eligió nuevamente Copilot. Dejemos estas diapositivas. Podemos dejar esta. Tiene un aspecto medio apocalíptico, ¿no? Hablemos entonces de la ofensa y la defensa en cuanto a IA. Quisiera comenzar con Jeff. ¿Qué prevé Jeff

en las próximas etapas y, cuando se fija en cómo está cambiando la IA, qué es lo que va a hacer con el área del uso indebido del DNS?

JEFF BEDSER: Qué buena pregunta, Jeff. Espero que no esté escuchando nadie. Creo que el espacio que vamos a ver con mayor presencia es en las identidades falsas, la posibilidad de la IA de crear prácticamente documentos de identidad casi perfectos con identidades de personas reales pero con fotos diversas, lo que imposibilite validar una persona real de una falsa.

Es una inquietud importante. Cuando pensamos que ahora podemos decir: "Bien, IA", no la de él, porque no quiero ser sarcástico. Quiero un número de tarjeta de crédito nuevo. Quiero que esté con una identidad, un domicilio en el estado de Pennsylvania en Estados Unidos. Quiero un registro de conductor conforme a toda esta información con mi fotografía y que no tenga puntos en contra. Quiero comprar un dominio con esa tarjeta de crédito inventada y pasaría todos los controles. Esto se ha generado en cuatro o cinco segundos.

Cuando hablamos de ciberdelito, normalmente esto se da con cuentas bancarias comprometidas, tarjetas de crédito robadas y tenemos credenciales básicamente robadas. Esas credenciales que validan una persona real son riesgos concretos y pueden cambiar rápidamente el panorama.

GREG SHATAN: Tim, le hago la misma pregunta pero también la pregunta previa. Respecto de los desarrolladores tempranos de IA, ¿qué están haciendo? ¿Son conscientes ellos del espacio del uso indebido del DNS en comparación con las otras cosas que puede utilizarse en la IA?

TIM MAURER: La respuesta es sí. Nosotros también tenemos una infraestructura de la nube importante. En general, la Internet trabaja de manera tal que se trata de evitar el uso indebido de los actores maliciosos. Hay algunas buenas noticias en esto. Desde la aparición de los distintos modelos que han sido implementados hay algunos modelos que implementaban un escenario más apocalíptico y no hemos visto el resurgimiento de algunas cosas que vimos al principio de los 2000 con la IA generativa que creaba nuevo código, que no permitía los mecanismos existentes de detección o el uso de filtros.

Lo que vimos es que se usa más IA para avanzar esos objetivos. Obviamente, esto no es bueno pero en términos del espectro de posibilidades creo que hay algunas buenas noticias en términos de lo que vemos. Creo que hay algunas lecciones interesantes mirando hacia los últimos meses respecto de cómo los utilizan los actores de defensa. Es divertido ver lo que se hace pero también resalta la importancia de cómo se utiliza la tecnología, lo que nos lleva a utilizarla, los prompts específicos y traducirlo en técnicas defensivas.

Si por ejemplo utilizamos Copilot en seguridad o algún otro producto o servicio, por supuesto para producir informes específicos cuando se detecta una anomalía utilizar para mayor automatización, mostrar una anomalía específica, elaborar un resumen ejecutivo y verificar de instruir al modelo también para errar por el lado de más en lugar de incluir menos anomalías potenciales para capturar los falsos positivos porque los humanos tienen que verificar todo esto y preferimos capturar más falsos positivos que podemos descartar, que se incluyan más falsos negativos.

El humano sigue siendo importante. La manera en que podemos aprovechar IA es utilizarla de manera tal que nos ayude a utilizar el ser humano de manera más efectiva. Hablando de escala, me hace ser optimista a pesar del consenso del que usted habló sobre la parte de identidad, con lo cual estoy totalmente de acuerdo.

Uno de los desafíos más importantes del lado de la defensa es la falta de escala y de gente capacitada para hacer determinadas cosas. Podemos lograr que la IA utilice esa escala reduciendo las barreras de entrada y para reducir las denuncias para que inclusive lo haga la gente que está menos capacitada. Esto mejoraría la evaluación del uso de IA en todo esto.

Respecto de lo que decía Ana, la importancia de participación entre el sector público y el privado, creo que mirando los últimos años es fascinante ver cómo después de ChatGPT y el New York Times en la página de apertura y ver también la ley de IA respecto de las cosas

que se realizaron, hubo una implementación muy rápida de nueva normativa y reglas y políticas en base a lo cual algunas empresas tuvieron que ver que se ajustaran a toda la normativa con la emergencia también de institutos, defensa contra las amenazas y también tener en cuenta el tema de los números de tarjeta de crédito. Vi la rapidez con la que los gobiernos pusieron una estructura muy rápidamente participando con el sector privado, Microsoft y otras empresas, para tratar de elaborar alguna infraestructura en base al riesgo para ver lo que hacía falta hacer.

GREG SHATAN: Gracias. Ana, me gustaría saber su perspectiva y particularmente la perspectiva europea, porque nosotros tres estamos en Estados Unidos constantemente que, por supuesto, está cambiando de manera importante pero me interesaría ver perspectivas diversas respecto de esta mezcla.

ANA NEVES: En Europa, creo que entendemos que respecto del uso indebido del DNS hay una evolución importante con la IA, respecto del aprendizaje por máquinas y la IA. Estamos tratando de desarrollar guías para el uso de IA y la protección del DNS considerando temas de privacidad respecto del análisis del DNS y la IA.

Tenemos un análisis realizado en la Unión Europea con un estudio de uso indebido del DNS y recomendaciones para los registros y registradores. Creo que desde diciembre de 2022 está vigente. Es

interesante ver dominios como .DE y .EU, que son menos abusivos, donde el uso indebido del DNS no es tan fuerte.

Esto prueba que la implementación técnica y el trabajo conjunto entre los sectores público y privado es algo muy importante. Trabajando juntos con los investigadores en IA, dado que son los que pueden ayudar en la prevención. Teniendo este componente de prevención que se adelante a lo que puede suceder es muy importante. Por ende, el trabajo de múltiples partes interesadas es muy importante en este proceso y el uso indebido del DNS es parte de todo esto.

La importancia de incluir investigadores y desarrolladores de la academia en todos estos procesos es fundamental. Con esta participación y escuchando todas las voces, las distintas partes interesadas, podemos lograr distintas capas en cuanto a política, ya sea a nivel público o privado en pos de este objetivo. Muchas gracias.

GREG SHATAN: Gracias, Ana. Tenemos muchas personas con preguntas, tanto aquí en el salón como remotamente. Si están participando remotamente, por favor, pongan sus preguntas en la caja de preguntas y respuestas. Michelle leerá las preguntas. Si están en el salón y en Zoom, pueden alzar la mano y hacer la pregunta. Si no están en Zoom, simplemente alcen la mano. Si no están en la mesa tenemos un micrófono que les podemos entregar. No sientan

timidez si no tienen un micrófono a su lado. Vamos a comenzar con Joanna Kulesza.

JOANNA KULESZA: Gracias. Gracias por sus ponencias. Es bueno escuchar el enlace entre IA y DNS. Tengo una pregunta que tiene que ver con la intervención de Ana y su último comentario. Creo que tal vez tenga sentido que intente decir esto. Gracias por su revisión integral. No sé si se siente cómoda contestando esto con su sombrero de Microsoft pero tenemos un trasfondo compartido de investigación. Por eso si quiere hablar de políticas está bien pero voy a comenzar con el trasfondo con respecto a la pregunta pero quiero su comentario o reacción.

Por un lado Microsoft fue vital para apoyar a Ucrania después de la agresión inesperada. Esto fue necesario para que Ucrania preparase sus defensas, especialmente ante las amenazas. Parece que tal vez la situación va a cambiar, parece que la información se robó pero se le va a devolver. Esto es una indicación clara de cómo regulaciones domésticas pueden atacar esto, especialmente en el nivel del uso indebido del DNS y el intercambio de información.

Por otro lado, Ana mencionó que existen reglamentos aumentados con respecto a IA. El Senador Cruz en diciembre del año pasado dijo que la ley de IA roba la innovación y puede considerarse como una interferencia extranjera en el desarrollo de la IA en los Estados Unidos.

Esto me lleva a mi última pregunta. Una buena amiga, Marietje Schaake, publicó un libro que se llama en inglés “The IT Coup”. Recientemente habló de ello. Ella mencionó que este podría ser el momento cuántico para los investigadores, quienes no se sienten muy cómodos al hacer sus investigaciones en los Estados Unidos, que se muden a la Unión Europea.

Ellos deberían darles la bienvenida completamente. Similar a lo que ocurrió en 1930. Con su sombrero de investigación, porque quizá no sea una pregunta cómoda para contestar de parte de Microsoft. Qué piensa usted en cuanto al desarrollo de la IA con respecto a ciberseguridad y qué puede hacer la comunidad de usuarios finales, tal vez trabajando en conjunto con el gobierno para asegurarnos de utilizar la IA para proteger a los individuos de estos ciberdelincuentes. Creo que esta es una forma diplomática de presentar esta pregunta. Creo que usted puede claramente identificar el problema geopolítico que estoy tratando, pero cualquier comentario que usted quiera compartir lo agradezco. Muchas gracias.

TIM MAURER: Gracias, Joanna. Qué bueno verla también. Esto tiene que ver con los días de prepandemia antes de estar con el gobierno pero puedo contestar esto bajo mi sombrera de Microsoft porque hemos sido claros en cuanto a estos retos geopolíticos y se compromete con la ciberseguridad para la empresa, para los clientes. Existe la iniciativa de un futuro seguro. Eso está por delante.

Quiero hablar acerca del último punto, en cuanto a qué puede hacer esta comunidad. Yo creo que nosotros sí estamos en un momento de eje, por decirlo así. Ha existido algún progreso y Microsoft ha lanzado productos y servicios nuevos, y lo hace regularmente y lo continuará haciendo. Es para ayudar a esta comunidad, cómo mejor utilizar la tecnología dentro de este contexto. Depende también de las personas que se enfocan en monetizar esto rápidamente. Ellos quieren probar estos productos para ver si pueden aprovecharse de ello.

Si identifican áreas donde se ven vulnerabilidades y donde consideran que hay preguntas de investigación donde Microsoft debería enfocarse, por favor, compartan eso con nosotros porque estamos trabajando en conjunto para poder maximizar la potencialidad de estos productos. Podemos trabajar con el GAC y con otras aportaciones de otras comunidades y de ustedes.

En cuanto al primer punto, en cuanto a la Unión Europea y los reglamentos, la empresa ha sido muy clara. Hay ciertas áreas donde los reglamentos serían útiles. Creo que la empresa y Microsoft empieza a preocuparse si esos reglamentos no se enfocan en los resultados, ya que la tecnología va a un ritmo veloz. Quizá vemos estos marcos regulatorios que surgen en estas jurisdicciones que quizá no está en unión con otros. Llega a ser un reto para nosotros navegar en ello e implementarlo, por eso pensamos que estamos en la cúspide de un avance donde tal vez

podamos a un nivel sistémico ponernos a favor de la defensa y queremos lograr eso.

GREG SHATAN: Gracias, Tim. Ahora sí, podemos ir a la caja de preguntas y respuestas. No sé quién está leyendo eso.

MICHELLE DESMYTER: Tenemos una pregunta de parte de [inaudible]. ¿Con qué método se encuentra la IA con CleanDNS?

JEFF BEDSER: Bueno, los métodos que utilizamos están en la presentación. Sé que internamente sí utilizamos Copilot para usos internos pero en la plataforma yo creo que eso es algo tipo propietario. Utilizamos algo disponible a nivel comercial pero no tengo la libertad de poder decir cuál es.

MICHELLE DESMYTER: La siguiente pregunta es de Sivasubramanian. Por otro lado, ¿acaso se puede utilizar IA para diseñar estos sitios de phishing para que se vean más auténticos hasta tal punto que los sitios de phishing estén libres de los patrones y los elementos que se ven típicamente en ello?

JEFF BEDSER: Puedo contestar a esa pregunta. Sí, pero un sitio de phishing no es solo la página de aterrizaje. Los detalles de esos sitios de phishing tienen que ver con lo siguiente. Hay sitios de phishing. Se hacen con un propósito malicioso. La edad del dominio es un factor y también la estructura. También los servidores de nombres subyacentes, los IP, la reputación de eso. Tal vez no se dan cuenta de que las direcciones de IP y el alojamiento son bulletproof hosts. No inscriben nada. No es una buena práctica. Es como si estuviera facilitando actividad criminal. Tiene que haber más que la página de aterrizaje. Las páginas de aterrizaje están mejorando y cada vez usan mejor la IA.

GREG SHATAN: Ahora le doy la palabra a Hadia Elminiawi, que está en el salón. Hadia se está acercando al micrófono.

HADIA ELMINIAWI: Gracias, Greg. Esto es para los registros. Soy Hadia. Mi pregunta es para Ana. Habló acerca de la transparencia en la IA y mi pregunta para usted, en términos prácticos, qué significa la transparencia. ¿Significa esto que es una IA de código abierto? Por ejemplo, sabemos que la IA de código abierto todavía no le quita el misterio de cómo IA toma sus decisiones. Los sistemas de blackbox todavía permanecen como cajas negras, aun con el código abierto. En términos prácticos, ¿qué significa esto en realidad?

ANA NEVES: Gracias por esa pregunta. Muy interesante. Yo creo que cuando yo mencioné la transparencia de la IA se trataba de la transparencia y el rendir cuentas. Los modelos de IA se deben poder explicar. Permitir los servidores de DNS y los profesionales de ciberseguridad que puedan entender los procesos de la toma de decisiones. Si no entiende por qué está peleando, eso es difícil. Si se pueden explicar los modelos de IA llega a ser más fácil tratar el tema de dónde existe el error. Eso es todo. Gracias.

GREG SHATAN: Jonathan, ¿tiene una pregunta?

JONATHAN ZUCK: Sí, la tengo. Yo solo tengo mi propia IA, que la llamaría inteligencia antigua. Detecta phishing y otros temas. Parece que alguno de los correos electrónicos de phishing están conectados de alguna manera con entidades de marcas registradas o marcas comerciales, o nombres reconocidos. En algún lugar está incrustado un enlace a un sitio de pharming. Tengo curiosidad. Uno que entre en el correo electrónico, si puede detectar de dónde proviene y ver los enlaces incrustados en el correo electrónico y mirar a ver si el dueño de esa entidad es el dueño que debe ser, el dueño de donde proviene. Es un trabajo que hace la IA y no he visto algo similar a eso o que alguien haga algo similar a eso.

Por ejemplo, cuando uno manda un correo electrónico donde uno adjunta un archivo y le dice a uno: “Mire, se te olvidó adjuntar el

archivo de su correo electrónico". Parece ser como si leyeran el correo electrónico. "Parece que es de Norton". Los enlaces incrustados vienen de Bitly, así que uno tiene que tener cuidado.

TIM MAURER: Creo que esto tiene que ver con lo que hablamos al principio. Cuando hablamos de IA, mucho de esto no es nuevo y ya tenemos algunos sistemas ya montados donde, en base a algoritmos y a aprendizaje automático, se dan cuenta de si esto ocurre automáticamente. El punto es que es más limpio de lo que era antes. El potencial real que esperamos ver con la IA es poder conectar mejor todos estos puntos y entenderlo mejor, y reunir estos recursos para poder detectar esto. Si un sitio se registra donde hay preocupación en cuanto a la identidad debido a la longitud de la página o algo así. Ahí es donde existe el potencial, que sea algo más amigable para el usuario. Que ella sepa qué leer y cómo actuar.

GREG SHATAN: Ya nos pasamos del tiempo pero voy a aceptar una pregunta más. Amrita.

AMRITA CHOUDHURY: Muchas gracias, Greg. Tengo dos preguntas. Me gustaría entender de Tim, usted y Jeffrey, que hemos estado hablando acerca del uso indebido del idioma inglés y cómo IA lo puede utilizar o no utilizarlo

al grabarlo. Pero qué pasa con los que no hablan inglés. Es mayormente de quienes hablamos.

Hemos visto que este uso indebido del dominio está creciendo en todos los lugares. Especialmente al venir de países donde tienen usuarios nuevos y no tienen conocimiento o son analfabetos. ¿Qué estamos haciendo para esos usuarios finales con esos vídeos o alguna clase? Por ejemplo, Microsoft hace 10 años tenía cursos de datos para que aprendan acerca de datos. Qué se está haciendo en este aspecto porque hay personas que están perdiendo dinero con respecto a eso.

TIM MAURER: Microsoft con sus negocios por todo el mundo está lanzando productos, servicios y modelos en varios idiomas. Tratan de seguridad. No solo ocurren en inglés sino también en otros idiomas. Estamos muy comprometidos con asegurarnos de que la tecnología sea asequible a las personas por todo el mundo y también medidas de seguridad y mitigación, para que reflejen también los usuarios que lo utilizan.

Con respecto a la capacidad, desarrollo de capacidad y el entrenamiento. Por ejemplo, en Kenia, enfocándonos en la parte sur del mundo, queremos asegurarnos de que ellos supieran cómo utilizar la tecnología y también tenemos capacitación interna. Se nos alienta a que nosotros mismos lo utilicemos. No solo se

concentra en los individuos y las empresas. También que nosotros podamos lanzar esos productos para maximizar esto.

JEFF BEDSER: Gracias por mencionar eso. Las herramientas que se utilizan están en cualquier idioma, para cualquier análisis o texto. Por eso la IA es tan eficaz para nosotros. Ya no necesito a alguien en mi personal que pueda leer o escribir todos los idiomas sino que la IA puede detectar cuál es el idioma y sacar los componentes relevantes. Es un componente grande el asegurarnos de que podamos hacer esto con todos los idiomas, porque estamos proporcionando servicios por todo el mundo. Vemos phishing, malware, en todos los idiomas. Utilizamos la IA eficazmente. Yo no podría costear un personal tan grande en todos los idiomas.

GREG SHATAN: Gracias, Jeff. Se nos acabó el tiempo. Dana, [inaudible], disculpen que no tenemos más tiempo para preguntas pero gracias a nuestros panelistas Jeff Bedser, Tim Maurer, Ana Neves, por estar aquí. Gracias por escucharnos y por permitir que esta mesa redonda de NARALO sea algo inolvidable. Volveremos para NARALO Townhall. Gracias. Terminamos la reunión.

[FIN DE LA TRANSCRIPCIÓN]