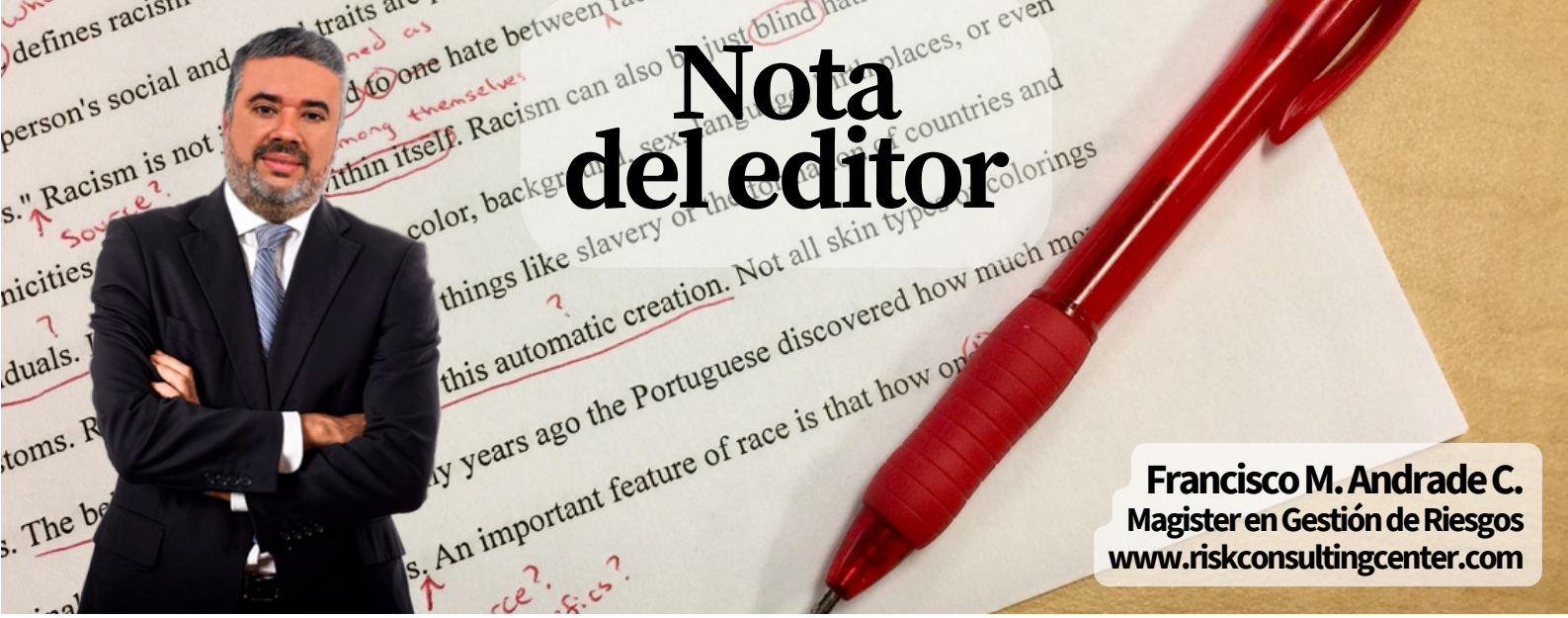
N. 8 | ENERO 2024

Hablemos de RIESGOS

ARTÍCULOS DE OPINIÓN

El DEEPFAKE, una nueva amenaza que incrementa el RIESGO REPUTACIONAL

Realizado por:
Francisco M. Andrade C.
Magister en Gestión de Riesgos
www.riskconsultingcenter.com



¡Cordial saludo estimado lector!

En esta ocasión, considero oportuno abordar un tema muy apremiante en esta era digital: los deepfakes y su relación con el riesgo reputacional. Este artículo no solo invita a la lectura, sino que también busca motivar a todos a tomar conciencia de la importancia de esta amenaza y a unirse en la búsqueda de soluciones efectivas.

En un mundo donde la información fluye a través de la red a una velocidad vertiginosa, la capacidad de la inteligencia artificial para manipular imágenes y audio plantea un desafío significativo. Los deepfakes, impulsados por algoritmos de aprendizaje profundo, tienen la capacidad de hacer que una persona aparezca diciendo o haciendo cosas que nunca hizo. Esta es una tecnología que ha evolucionado rápidamente en los últimos años.

¿Por qué debería preocuparnos? La respuesta es simple: los deepfakes tienen el potencial de destruir reputaciones, influir en procesos democráticos, difamar a individuos y empresas, y crear un clima de desconfianza generalizada. Los casos recientes de deepfakes utilizados para desinformación política, difamación personal y fraudes empresariales ilustran la urgente necesidad de abordar esta amenaza de manera efectiva.

Este artículo de opinión no solo explora los casos más impactantes y las características alarmantes de los deepfakes, sino que también presenta medidas de seguridad y control que pueden ayudarnos a protegernos contra esta amenaza.

Desde la detección automatizada hasta la educación pública, y la colaboración entre plataformas digitales, hay pasos concretos que podemos tomar para enfrentar esta creciente preocupación.

La lectura de este artículo no solo es una oportunidad para informarse sobre un tema de relevancia actual, sino que también es un llamado a la acción. Todos nosotros, como ciudadanos, consumidores de medios y usuarios de tecnología, debemos tomar medidas para proteger nuestra reputación y la integridad de la información que consumimos y compartimos en línea.

Los deepfakes representan un desafío que trasciende las fronteras y las industrias, y requiere la cooperación de la sociedad, la industria y los gobiernos para abordarlo de manera efectiva. Invitamos a todos a leer este artículo con atención, a compartirlo ampliamente, y a unirse en la búsqueda de soluciones para proteger nuestra integridad personal y empresarial en la era digital.

La amenaza de los deepfakes es real, pero juntos podemos enfrentarla y garantizar que la verdad y la confianza prevalezcan en nuestro mundo digital.

En la era digital actual, la inteligencia artificial (IA) ha avanzado a pasos agigantados y ha revolucionado diversos aspectos de nuestra vida cotidiana. Uno de los aspectos más preocupantes y controvertidos de esta revolución es la proliferación de los deepfakes. Los deepfakes son manipulaciones audiovisuales generadas por IA que pueden hacer que una persona parezca decir o hacer cosas que nunca hizo. A medida que esta tecnología continúa evolucionando, es fundamental abordar sus riesgos, características y buscar medidas de seguridad y control.

Los deepfakes son el resultado de la intersección de la IA y la manipulación de medios digitales. La tecnología detrás de los deepfakes se basa en algoritmos de aprendizaje profundo que pueden analizar y sintetizar grandes cantidades de datos para crear videos o grabaciones de audio falsificados con una precisión sorprendente. Estos algoritmos han evolucionado rápidamente en los últimos años, gracias a la disponibilidad de grandes conjuntos de datos y el aumento de la capacidad computacional.

Una de las características más notables de los deepfakes es su capacidad para imitar las voces y los gestos faciales de manera realista, lo que hace que sea extremadamente difícil distinguirlos de contenido genuino. Además, los deepfakes pueden ser generados en poco tiempo y con recursos accesibles, lo que ha permitido que cualquier persona con conocimientos básicos de IA pueda crearlos. Esto ha llevado a la aparición de una gran cantidad de deepfakes en línea, que van desde bromas inofensivas hasta contenido dañino y manipulativo.

Según el informe ***"El Deepfake que Genera la IA: Este es el Nuevo Riesgo que no Debe Pasar por Alto el Seguro"***, publicado en la página web de Future de Inese, este aborda las implicaciones y riesgos asociados con la tecnología de deepfake asi:

1. Aumento Exponencial de Fraudes con Deepfake:

- Según un informe de Onfido, los intentos de realizar estafas utilizando tecnología deepfake aumentaron un 3000% en 2023. Esta drástica alza se atribuye a la creciente disponibilidad de herramientas en línea económicas y sencillas, así como al uso de Inteligencia Artificial generativa.

- Las aplicaciones de intercambio de caras (face-swapping) son comunes en deepfakes. Estas pueden ser desde pegar burdamente una cara sobre otra hasta utilizar IA para transformar y combinar rostros de manera más realista. Los estafadores pueden hacerse pasar por otras personas, incluyendo conocidos o celebridades.

2. Fraudes Menos Sofisticados:

- Los programas informáticos que utilizan esta tecnología son fáciles de ejecutar y económicos, algunos incluso gratuitos. En muchos casos, una misma serie de falsificaciones puede utilizarse simultáneamente en múltiples ataques.

- Estas falsificaciones menos sofisticadas, conocidas como "cheapfakes", pueden intentar aplicarse en sistemas de verificación facial, realizar transacciones fraudulentas o acceder a información comercial sensible. Aunque la mayoría son evidentes, los criminales las utilizan en amplios ataques, esperando que alguno tenga éxito.

Casos Recientes

Para comprender la amenaza que representan los deepfakes, es importante revisar algunos casos recientes que han sacudido el paisaje mediático y político.

Política y Desinformación: Durante las elecciones presidenciales de Estados Unidos en 2020, surgieron varios deepfakes que pretendían mostrar a los candidatos en situaciones comprometedoras o haciendo declaraciones falsas. Estos videos falsos pueden influir en la opinión pública y socavar la confianza en el proceso electoral.

Crimen y Difamación: Los deepfakes también se han utilizado con fines delictivos, como la difamación de personas inocentes. Por ejemplo, se han creado videos pornográficos falsos con la imagen de celebridades y personas comunes, lo que constituye una invasión flagrante de la privacidad y puede causar un daño irreparable a la reputación de las víctimas.

Fraude Empresarial: En el ámbito empresarial, los deepfakes pueden utilizarse para engañar a empleados y socios comerciales. Los estafadores han suplantado la voz de CEOs para autorizar transacciones financieras fraudulentas, lo que resulta en pérdidas económicas y un daño significativo a la reputación de la empresa.

Campañas de Odio: Los deepfakes también se han utilizado en campañas de odio y discriminación. Se han creado videos falsos que difaman a grupos étnicos, religiosos o políticos, lo que puede tener consecuencias graves en términos de imagen y reputación.

Características de los Deepfakes

Para entender mejor la amenaza que representan los deepfakes, es esencial conocer algunas de sus características más relevantes:

1. Realismo Asombroso: Los deepfakes pueden crear videos y grabaciones de audio que son casi indistinguibles de los originales. Esto hace que sea extremadamente difícil detectarlos a simple vista u oído.

2. Facilidad de Creación: La disponibilidad de software y herramientas de generación de deepfakes ha democratizado su creación. Cualquier persona con acceso a Internet y tutoriales en línea puede generar deepfakes.

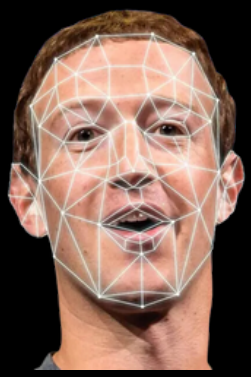
3. Difusión Rápida: Las redes sociales y plataformas de intercambio de videos permiten que los deepfakes se propaguen rápidamente, alcanzando audiencias globales en cuestión de horas.

4. Manipulación de la Realidad: Los deepfakes pueden alterar la realidad de manera significativa, haciendo que sea difícil distinguir entre lo auténtico y lo falso. Esto socava la confianza en los medios de comunicación y la información en general.

5. Escalada de Conflictos: En el ámbito internacional, los deepfakes podrían utilizarse para crear conflictos y tensiones, ya que pueden hacer que líderes mundiales aparenten declaraciones o acciones provocativas.

Casos Impactantes de Deepfakes

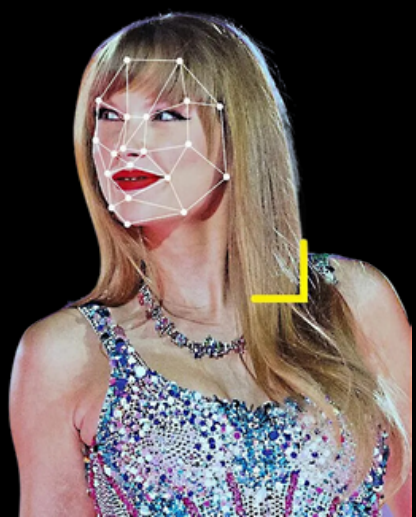
Barack Obama: En 2018, se creó un deepfake del expresidente de los Estados Unidos en el que parecía decir cosas que nunca dijo. Este video se viralizó y puso de manifiesto la capacidad de los deepfakes para manipular la percepción pública de los eventos políticos.



Mark Zuckerberg: En 2019, un deepfake del CEO de Facebook, apareció en línea. En el video, parecía confesar secretamente el control de los datos de usuarios. Aunque fue creado como una demostración de la tecnología, planteó preocupaciones sobre la manipulación de figuras públicas y su impacto en la reputación de las empresas.

Tom Cruise: En 2021, se viralizó un video deepfake en el que un usuario imitaba perfectamente a Tom Cruise. Esto planteó preguntas sobre la autenticidad de los contenidos en línea y cómo podrían afectar la percepción de celebridades y su reputación.

Taylor Swift: 2024, la famosa cantante, se ha convertido en la nueva víctima de la IA debido a la proliferación de deepfakes explícitos sobre ella. Estas imágenes han inundado las redes sociales, incluyendo la plataforma X, donde algunas publicaciones han alcanzado hasta 45 millones de visitas. En este caso, han sido utilizados para crear contenido sexualizado.



Medidas de Seguridad y Control de Riesgos

Ante la creciente amenaza de los deepfakes y su conexión con el riesgo reputacional, es imperativo implementar medidas de seguridad y control efectivas. A continuación, se presentan algunas recomendaciones clave:.

1. Capacitación y Concientización del Personal:

- Implementar programas de capacitación que eduquen a los empleados sobre los riesgos asociados con los deepfakes y cómo identificar posibles intentos de fraude o manipulación.
- Fomentar una cultura de conciencia digital y ética, donde los empleados comprendan la importancia de proteger la integridad de la empresa frente a amenazas de reputación.

2. Protocolos de Verificación:

- Establecer protocolos rigurosos para verificar la identidad de personas que realizan transacciones financieras o autorizan decisiones críticas mediante comunicación remota.
- Utilizar métodos de autenticación multifactorial para garantizar la legitimidad de las interacciones, especialmente cuando se trata de información sensible.

3. Auditorías de Seguridad y Riesgos:

- Realizar auditorías regulares de seguridad y riesgos para evaluar la efectividad de las medidas de mitigación y prevención implementadas.
- Contratar servicios de terceros especializados en seguridad informática para identificar posibles vulnerabilidades y mejorar continuamente las defensas de la empresa contra deepfakes y otras amenazas cibernéticas.

4. Monitoreo Proactivo:

- Establecer equipos dedicados para monitorear activamente las redes sociales y otras plataformas digitales en busca de deepfakes que puedan afectar la reputación de la empresa.
- Responder rápidamente a cualquier intento de difamación o desinformación, y colaborar con las plataformas para eliminar contenido falso de manera expedita.

5. Implementación de Marcas de Agua Digitales:

- Explorar la incorporación de marcas de agua digitales en los medios visuales producidos por la empresa, lo que facilitaría la verificación de la autenticidad y actuaría como un disuasivo adicional contra manipulaciones.

6. Contratos y Acuerdos de Confidencialidad:

- Revisar y fortalecer los contratos y acuerdos de confidencialidad con socios comerciales y proveedores, especificando claramente las medidas de seguridad que se esperan y las consecuencias en caso de incumplimiento.

7. Inversión en Tecnologías de Detección Avanzada:

- Explorar y adoptar tecnologías avanzadas de detección de deepfakes, como sistemas de aprendizaje automático que analizan patrones sutiles para identificar manipulaciones en contenido multimedia.

8. Participación Activa en Iniciativas de Industria:

- Colaborar con otras empresas del sector para compartir información sobre amenazas y mejores prácticas en la prevención de deepfakes y otros riesgos reputacionales.
- Participar en iniciativas de la industria que buscan establecer estándares y protocolos para proteger la integridad digital y la reputación empresarial.

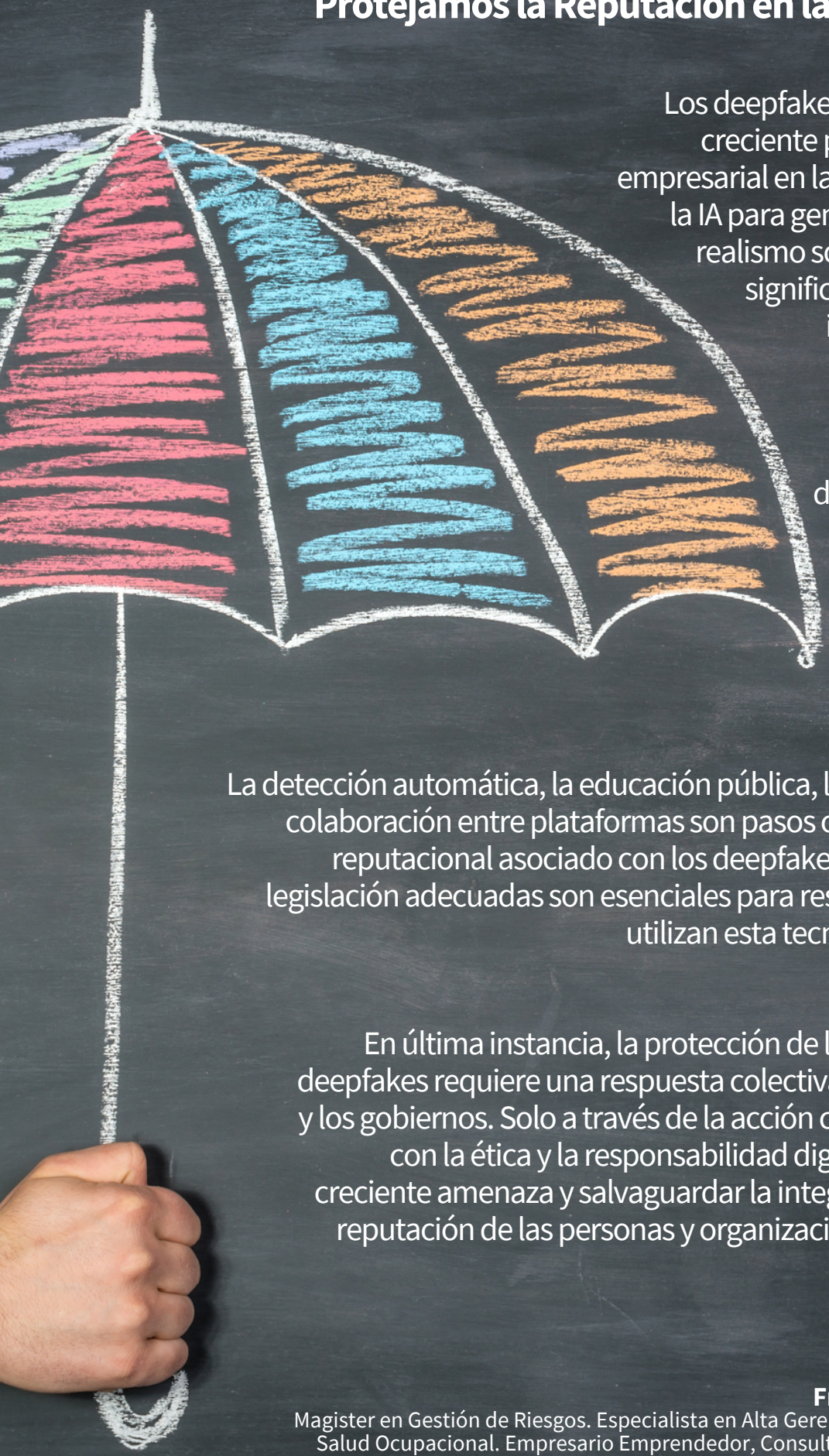
9. Revisión y Actualización Periódica de Políticas de Seguridad:

- Mantener políticas de seguridad actualizadas y relevantes para hacer frente a las evoluciones tecnológicas y amenazas emergentes, incluyendo orientaciones específicas sobre la manipulación de medios digitales.

10. Promover la Transparencia y la Comunicación Proactiva:

- Ser proactivos en la comunicación con stakeholders en situaciones en las que la empresa pueda estar expuesta a riesgos reputacionales.
- Proporcionar información transparente sobre las medidas de seguridad implementadas y los esfuerzos en curso para mitigar riesgos.

Al adoptar un enfoque holístico que combine tecnología avanzada, educación del personal y colaboración con la comunidad empresarial, las organizaciones pueden fortalecer sus defensas contra el riesgo reputacional asociado con los deepfakes. La prevención y la preparación son clave en el mundo digital en constante evolución.



Conclusiones:
Protejamos la Reputación en la Era de los Deepfakes

Los deepfakes representan una amenaza creciente para la integridad personal y empresarial en la era digital. La capacidad de la IA para generar contenido falso con un realismo sorprendente plantea riesgos significativos para la reputación de individuos y organizaciones.

Los casos recientes de desinformación política, difamación personal y fraude empresarial destacan la necesidad de abordar esta amenaza de manera urgente.

La detección automática, la educación pública, la verificación de fuentes y la colaboración entre plataformas son pasos críticos para mitigar el riesgo reputacional asociado con los deepfakes. Además, la regulación y la legislación adecuadas son esenciales para responsabilizar a aquellos que utilizan esta tecnología con fines maliciosos.

En última instancia, la protección de la reputación en la era de los deepfakes requiere una respuesta colectiva de la sociedad, la industria y los gobiernos. Solo a través de la acción coordinada y el compromiso con la ética y la responsabilidad digital podemos enfrentar esta creciente amenaza y salvaguardar la integridad de la información y la reputación de las personas y organizaciones en un mundo cada vez más digital y conectado.

El autor

Francisco M. Andrade C. MScGR

Magister en Gestión de Riesgos. Especialista en Alta Gerencia. Administrador de la Seguridad y Salud Ocupacional. Empresario Emprendedor, Consultor Empresarial, Auditor y Profesor en administración y gestión integral de riesgos, alta gerencia, gestión de riesgos en proyectos y logística integral, seguridad física empresarial y seguridad & salud laboral en instituciones de educación superior (Cursos, Diplomados, Especializaciones y Maestrías). Experiencia de más de 20 años en funciones directivas inherentes a la Seguridad Integral y de la Gestión de Riesgos Empresariales. Gerente General de Risk Consulting Center SAS.

Especialistas en Administración y Gestión Integral de Riesgos - RISK® CONSULTING CENTER

(057) 311 809 5576
BOGOTÁ - COLOMBIA
www.riskconsultingcenter.com

Website
RISK CONSULTING CENTER SAS

SÍGUENOS EN NUESTRAS REDES