Guía para la gobernanza: El uso de COSO 2013 en la era de la IA

Equilibrando el control y la innovación.

Junio 2025

¿Qué es COSO 2013?

- **Marco de confianza**

  El Marco Integrado de Control Interno COSO 2013 proporciona una estructura reconocida para evaluar y mejorar los controles internos.

- **Cinco componentes**

  El entorno de control, la evaluación de riesgos, las actividades de control, la información y comunicación, y las actividades de monitoreo forman la base.

- **17 principios**

  Estos principios guían a las organizaciones en la implementación de controles efectivos en los objetivos operativos, de información y de cumplimiento.

A medida que la IA transforma los procesos empresariales, las organizaciones deben asegurarse de que sus controles internos evolucionen en consecuencia. COSO 2013 proporciona la estructura perfecta para gobernar la IA de manera responsable.

Por qué la IA necesita gobernanza

La IA es cada vez más integral para las operaciones, la toma de decisiones y la estrategia en todos los sectores. Si bien ofrece oportunidades significativas, como eficiencia, conocimientos e innovación, también introduce nuevos riesgos:

Sesgo algorítmico

Los sistemas de IA pueden perpetuar o amplificar los sesgos existentes en los datos, lo que lleva a resultados injustos.

Vulnerabilidades de seguridad

Los sistemas de IA pueden ser susceptibles a la manipulación, el envenenamiento de datos o los ataques adversarios.

Riesgos de cumplimiento

La rápida evolución de las regulaciones en torno a la IA requiere una vigilancia y adaptación constantes.

01 Entorno de control: establecer el tono

El entorno de control establece los cimientos para la gobernanza de IA a través de los valores organizacionales, el liderazgo y la estructura.

Integridad y ética

Establecer principios y comités de ética de IA.

Supervisión del consejo

Asegurar que el liderazgo entienda y gobierne la IA.

Estructura organizacional

Definir roles y responsabilidades claros de IA.

Competencia

Desarrollar habilidades y conocimientos de IA.

Responsabilidad

Responsabilizar a las personas por los resultados de IA.

Supervisión del Directorio

Upskilling del Directorio

Proporcionar sesiones de educación regular para los miembros del Directorio sobre los riesgos emergentes de la IA, los desarrollos normativos y las principales iniciativas de IA.

Experiencia

Incluir miembros del Directorio o asesores con experiencia en tecnología y gobernanza de IA.

Informes de gobernanza

Establecer informes periódicos al Directorio sobre las actividades de IA, incluido un inventario de los principales sistemas de IA, sus propósitos, beneficios y riesgos asociados.

Sin una supervisión adecuada del Directorio, las iniciativas de IA pueden proceder sin una evaluación de riesgos o controles adecuados.

Responsabilidad por los resultados de la IA

Las organizaciones deben establecer una responsabilidad clara por los sistemas de IA para evitar una "brecha de responsabilidad" cuando algo sale mal.

RACI para IA

Desarrollar una matriz RACI (Responsable, Aprobador, Consultado e Informado) para los procesos clave de IA, asegurando que no haya ambigüedad sobre quién es responsable de qué.

Humano en el bucle

Para decisiones críticas, mantener la supervisión humana con una responsabilidad clara. Se debe establecer el principio de que la responsabilidad no se puede delegar a una máquina.

Responsabilidad por incidentes

Cuando ocurran incidentes relacionados con la IA, realizar revisiones posteriores al incidente que examinen si el personal no cumplió con sus responsabilidades.

02 Evaluación de riesgos: identificación de amenazas de IA

La evaluación de riesgos para la IA requiere un proceso dinámico e iterativo para identificar qué podría salir mal con los sistemas de IA.

Especificar objetivos

Definir objetivos claros y medibles para cada sistema de IA.

Identificar riesgos

Identificar sistemáticamente qué podría salir mal.

Evaluar el riesgo de fraude

Considerar cómo se podría explotar o manipular la IA.

Monitorear cambios

Identificar cambios que podrían afectar la efectividad de la IA.

Establecer objetivos claros de IA

Objetivos vagos como "queremos innovar con IA" son insuficientes. Las organizaciones deben especificar con claridad lo que cada sistema de IA debe lograr.

Objetivos SMART

Asegúrese de que los objetivos de IA sean Específicos, Medibles, Alcanzables, Relevantes y Temporales (SMART, su sigla en inglés). Por ejemplo: “mejorar la satisfacción con la respuesta del servicio al cliente en un 15% dentro de un año a través de un chatbot de IA".

Alineación con el apetito de riesgo

Al especificar los objetivos de IA, articule el apetito de riesgo relacionado con ese objetivo, como las tasas de error aceptables o los niveles de confianza.

Documentación y comunicación

Documente claramente los objetivos del sistema de IA y asegúrese de que todos los interesados los entiendan para evitar la falta de alineación entre los objetivos técnicos y empresariales.

03 Actividades de control: proteger a la IA

Las actividades de control son las acciones que ayudan a garantizar que se lleven a cabo las directivas de la gerencia para mitigar los riesgos. Para la IA, esto significa implementar controles específicos para asegurar que los sistemas operen dentro de los parámetros deseados.

Seleccionar controles

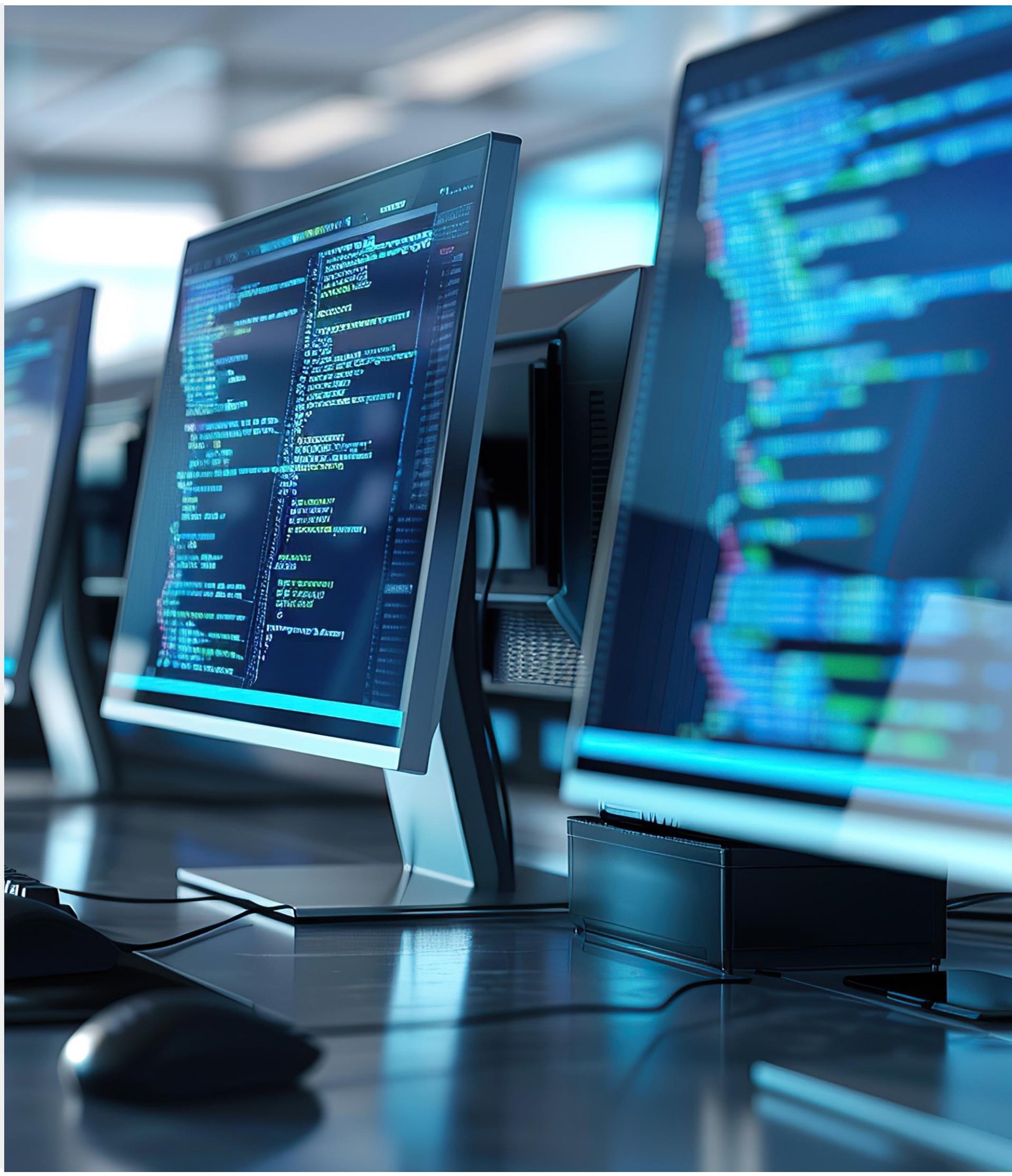
Elegir controles para mitigar los riesgos de la IA.

Controles generales de TI

Implementar salvaguardas tecnológicas.

Políticas y procedimientos

Documentar y comunicar las expectativas.



Actividades de control de IA esenciales

Validación previa a la implementación

Pruebas rigurosas antes de la implementación, incluida la validación en datos , pruebas de aceptación del usuario y pruebas de sesgo y equidad.

Límites y reglas automatizados

Superposición de controles basados en reglas para condiciones de límite, como apagados automáticos si un sistema de IA excede los umbrales predefinidos.

Monitoreo y revisión de resultados

Establecer controles de detección para revisar las decisiones/resultados de la IA, como hacer que los oficiales de cumplimiento revisen una muestra de las aprobaciones y denegaciones de la IA.

Redundancia y respaldos

Asegurar controles de respaldo y procedimientos manuales que puedan hacerse cargo si el sistema de IA falla o produce resultados cuestionables.

Controles generales de TI para IA

Los controles generales de TI (ITGC) garantizan que el entorno tecnológico que respalda la IA sea confiable y seguro.

Desarrollo seguro

Asegúrese de que el desarrollo del modelo de IA siga prácticas de codificación segura en entornos segregados antes de la promoción a producción.

Gestión de cambios

Implemente un control de cambios formal para los modelos de IA, incluida la documentación, la aprobación y las pruebas para cualquier cambio.

Controles de acceso

Limite el acceso a los sistemas y datos de IA mediante el control de acceso basado en roles, asegurando que solo el personal autorizado pueda acceder a los conjuntos de datos de entrenamiento o implementar modelos.

Resiliencia y recuperación

Incluya los sistemas de IA en los planes de recuperación de desastres, con copias de seguridad regulares del código del modelo y los estados del modelo entrenado.

Políticas y procedimientos para IA

Desarrollar una política de gobernanza de IA

Crear una política general que cubra los principios para el uso de IA, roles y responsabilidades, y referencias a procedimientos específicos.

Procedimientos de control de documentos

Para cada sistema de IA importante, crear procedimientos paso a paso para la operación, el manejo de excepciones y las vías de escalamiento.

Establecer procedimientos del ciclo de vida del modelo

Institucionalizar el ciclo de vida del modelo de IA con procedimientos para el desarrollo, la validación, la implementación y el monitoreo.

Crear planes de respuesta a incidentes

Desarrollar procedimientos para cuando los sistemas de IA produzcan errores o se descubran brechas/errores.

Capacitar y comunicar

Asegurarse de que todo el personal relevante entienda los procedimientos a través de capacitación, simulacros y comunicación regular.

04 Información y comunicación: la sangre vital

La información de calidad y los canales de comunicación claros son esenciales para una gobernanza eficaz de la IA.

Información de calidad

Asegúrese de que los sistemas de IA utilicen datos precisos y completes.

- Gobierno de datos
- Métricas de calidad
- Controles de validación

Comunicación interna

Comparta el conocimiento de IA dentro de la organización.

- Mensajes de liderazgo
- Reuniones interdepartamentales
- Capacitación y sensibilización

Comunicación externa

Involucre a las partes interesadas externas.

- Transparencia con los clientes
- Participación regulatoria
- Gestión de proveedores

El papel fundamental de la calidad de los datos

Para los sistemas de IA, los datos son la sangre vital. Si los datos que entran son deficientes, los resultados de la IA también lo serán: "basura entra, basura sale".

Marco de gobernanza de datos

Establecer un programa que defina los datos críticos para la IA y controle su calidad a través de diccionarios de datos, seguimiento de linaje y métricas de calidad.

Controles de validación de datos

Implementar controles automatizados y manuales para validar los datos antes de que los use la IA, desde la validación de entrada simple hasta *scripts* de conciliación complejos.

Monitoreo continuo de datos

Para la IA en producción, monitorear los datos entrantes en busca de cambios o anomalías que puedan indicar cambios en los patrones subyacentes o problemas en el *pipeline* de datos.

Sesgo en el reconocimiento facial: una lección sobre la calidad de los datos

Un ejemplo típico de problemas de calidad de datos que afectan a la IA es el caso de las tecnologías de reconocimiento facial que tienen tasas de error más altas para ciertos grupos demográficos.

Causa raíz

Los datos de entrenamiento no eran lo suficientemente representativos, con una proporción mucho mayor de rostros de hombres de piel clara y menos rostros de mujeres de piel oscura.

Impacto en la calidad

La salida de la IA (coincidencias faciales) tenía baja calidad para algunos grupos, lo que provocaba tasas de error más altas y posible discriminación.

Mejora

El reconocimiento de este problema ha llevado a las empresas a mejorar drásticamente la calidad de los conjuntos de datos de imágenes faciales mediante la recopilación de imágenes más diversas.

Comunicación interna efectiva

La comunicación interna asegura que todos entiendan los sistemas de IA, sus objetivos y las responsabilidades de control.

Mensajes de la gerencia

El liderazgo debe comunicar sobre las iniciativas de IA, su importancia para los objetivos y las responsabilidades de los empleados. Esto establece el tono y aclara las expectativas.

Colaboración entre funciones

Establezca reuniones periódicas donde diferentes partes interesadas (TI, ciencia de datos, unidades de negocio, riesgo, cumplimiento, auditoría) puedan compartir conocimientos y eliminar los silos.

Comunicación ascendente

Anime a los empleados a informar sobre problemas o sugerencias relacionados con la IA a través de canales claros, asegurándose de que la gerencia esté receptiva a las preocupaciones.

Estrategias de comunicación externa

Transparencia con el cliente

Si se utiliza IA en formas que interactúan con los clientes, comuníquelo abiertamente. Agregue divulgaciones como “esta recomendación fue generada por un algoritmo" o etiquete claramente las interacciones con chatbots.

Participación regulatoria

Mantenga una comunicación activa con los reguladores sobre sus implementaciones de IA. Comparta los resultados de validación o metodologías e infórmeles de inmediato sobre cualquier problema significativo.

Relaciones públicas

Tenga un plan para la comunicación externa en caso de que un incidente de IA se haga público. Es mejor que las partes interesadas escuchen sobre los problemas de usted primero, con su explicación y plan de remediación.

05 Actividades de monitoreo: vigilancia continua

El monitoreo asegura que los controles alrededor de la IA y los sistemas de IA mismos continúen operando de manera efectiva con el tiempo.

Evaluaciones continuas

Monitoreo continuo integrado en los procesos para verificar el desempeño de la IA.

Evaluaciones separadas

Auditorías o evaluaciones periódicas para examinar a fondo los controles de la IA.

Comunicación de deficiencias

Informar oportunamente sobre los problemas a quienes pueden solucionarlos.

Monitoreo continuo del rendimiento de IA

Los sistemas de IA no son "configurar y olvidar". Requieren un monitoreo continuo para detectar problemas antes de que causen daños.

Paneles de rendimiento

Implementar paneles que rastreen la precisión de las predicciones, las tasas de error y la distribución de la salida.

Alertas de umbral

Definir umbrales que desencadenen alertas o revisiones. Por ejemplo, si la precisión mensual de una IA cae por debajo del 90%, marcarla para investigación.

Verificaciones de efectividad de control

Usar tecnología para verificar continuamente que las actividades de control se están llevando a cabo según lo requerido por la política, como asegurar que se obtengan aprobaciones para las decisiones de IA de alto valor.

Evaluación de sistemas de IA

Auditorías internas

Asegúrese de que la auditoría interna incluya los sistemas de IA en su plan, evaluando no solo los controles de TI, sino también la gestión del riesgo del modelo, el cumplimiento ético y la calidad de los datos.

Comités de revisión de modelos

Establezca comités que revisen periódicamente todos los modelos significativos, examinando el rendimiento reciente, los incidentes y los cambios planificados.

Evaluaciones externas

Para sistemas de IA críticos o sensibles, considere evaluaciones externas como pruebas de penetración, auditorías de sesgo por parte de especialistas de terceros o certificaciones formales.

Estas evaluaciones separadas complementan el monitoreo continuo al proporcionar una evaluación más profunda e independiente de los controles y el rendimiento de la IA.

Comunicación de deficiencias de IA

Cuando se identifican fallas de control o problemas con la IA, se deben informar de inmediato a quienes puedan abordarlos.

Definir “deficiencia"

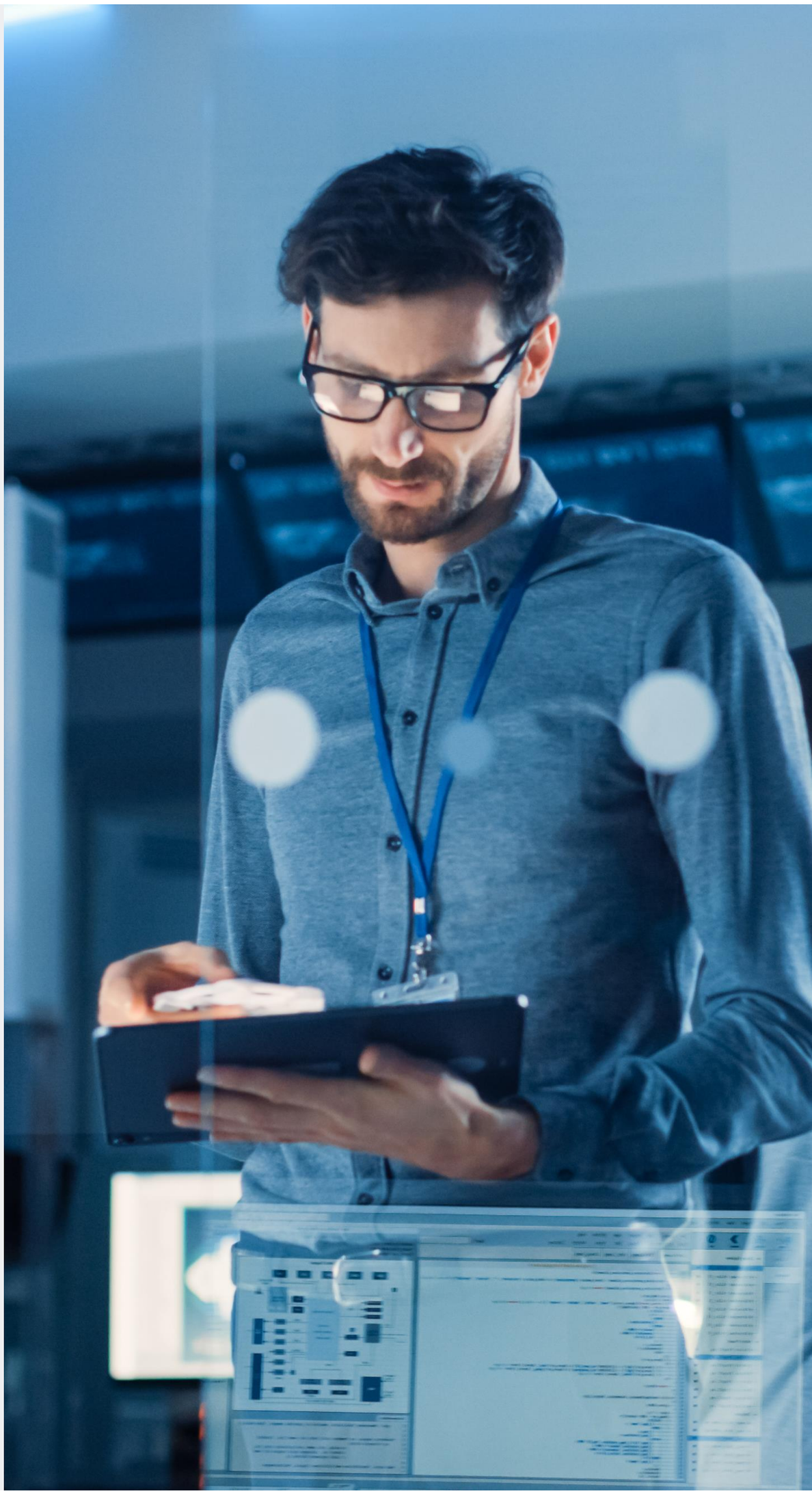
Establecer claramente lo que constituye una deficiencia en los controles o el rendimiento de la IA, como fallas para detectar ciertos eventos, omisión de pasos de validación o violaciones éticas.

Proceso de reporte de incidentes

Integrar los incidentes de IA en el sistema de gestión de problemas existente, asegurando la notificación automática a los propietarios de control relevantes y las vías de escalamiento apropiadas.

Informe al Directorio

Proporcionar informes periódicos al Directorio sobre las deficiencias de control interno que incluyan asuntos relacionados con la IA, con escalamiento inmediato para problemas graves.



El futuro de la gobernanza de IA

A medida que las tecnologías y las regulaciones de IA evolucionan, las organizaciones deben adaptar continuamente sus marcos de control internos.

Regulaciones emergentes

Leyes como la Ley de IA de la Unión Europea (UE) requerirán una gestión de riesgos formal y supervisión para aplicaciones de IA de alto riesgo, lo que hace que la gobernanza alineada con COSO sea esencial.

Evolución de la auditoría de IA

Están surgiendo nuevas metodologías para "auditar algoritmos", con herramientas y técnicas especializadas para evaluar la equidad, la transparencia y la seguridad de la IA.

Gobernanza integrada

Las organizaciones líderes están pasando de proyectos de IA aislados a marcos de gobernanza de IA a nivel empresarial que se alinean con las estructuras de riesgo y control existentes.

Tomar acción:
implementar COSO para IA hoy

La inteligencia artificial no es ni una bala mágica ni una fuerza incontrolable, es una herramienta poderosa que debe implementarse con rigor y supervisión.

Al mapear la gobernanza de IA al Marco COSO 2013, su organización puede innovar con confianza mientras mantiene el control. Comience evaluando su gobernanza de IA actual en relación con los cinco componentes de COSO, identificando brechas y desarrollando un plan de acción.

Recuerde:
Las organizaciones que gobiernan la IA de manera efectiva obtendrán una ventaja competitiva a través de la innovación y la confianza.

Contacto

Gonzalo Riederer

Socio Risk Assurance

PwC Chile

gonzalo.riederer@pwc.com

+56 9 8818 4606