



Consejo Nacional de Promoción y Apoyo a la Micro,
Pequeña y Mediana Empresa

MATERIAL DE ESTUDIO

Módulo 6 Uso seguro, responsable y evaluación

Protección de datos, verificación de resultados y supervisión humana

Curso	Inteligencia Artificial Aplicada a las MIPYMES
Duración del módulo	3 horas
Público	Emprendedores y propietarios de MIPYMES

Programa de formación para emprendedores

ORIENTACIONES

Propósito y resultados de aprendizaje

Este módulo presenta criterios prácticos para utilizar la inteligencia artificial de forma segura y responsable en una MIPYME. El propósito es reconocer riesgos, proteger la información, verificar los resultados y mantener la responsabilidad humana antes de utilizar una respuesta generada por IA.

Durante el estudio el participante aprenderá a distinguir información que puede compartirse de datos que deben protegerse, identificar señales de una respuesta poco confiable, revisar posibles sesgos, documentar decisiones y aplicar una lista de evaluación antes de incorporar un resultado a una actividad empresarial.

Al finalizar este estudio podrá

- Reconocer riesgos frecuentes asociados al uso de inteligencia artificial en una MIPYME.
- Proteger datos personales, financieros, comerciales y demás información sensible.
- Verificar respuestas para detectar errores, omisiones, supuestos e información inventada.
- Identificar situaciones que requieren transparencia, revisión de sesgos y supervisión humana.
- Aplicar criterios para decidir si un resultado puede utilizarse, corregirse o descartarse.

Ruta de estudio

TEMA	PREGUNTA ORIENTADORA
Uso responsable	¿Qué riesgos deben considerarse antes de utilizar una IA?
Protección de datos	¿Qué información puede compartirse y cuál debe protegerse?
Verificación	¿Cómo detectar errores, omisiones o información inventada?
Sesgos y transparencia	¿Cómo evitar resultados injustos o comunicaciones engañosas?
Supervisión humana	¿Qué decisiones requieren control y autorización de una persona?
Evaluación	¿Cómo decidir si un resultado es seguro, útil y apto para el negocio?

TEMA 1

Uso seguro y responsable de la inteligencia artificial

Utilizar inteligencia artificial de forma responsable significa aprovechar sus beneficios sin delegar de manera automática decisiones, datos o compromisos que puedan afectar al negocio, a sus clientes o a otras personas. La rapidez de una respuesta no garantiza que sea correcta, apropiada o segura.

Antes de utilizar una herramienta conviene preguntarse qué información necesita, qué consecuencias podría tener un error, quién revisará el resultado y si la tarea implica dinero, datos personales, derechos, seguridad o decisiones sensibles. Mientras mayor sea el impacto posible, mayor debe ser la supervisión.

Riesgos frecuentes y controles básicos

RIESGO	EJEMPLO	CONTROL RECOMENDADO
Información incorrecta	La IA genera un precio, fecha o dato no confirmado.	Verificar con la fuente original antes de usarlo.
Uso inadecuado	Se delega una decisión que requiere criterio humano.	Definir límites y aprobación obligatoria.

Principio esencial

La IA debe apoyar el trabajo, no sustituir la responsabilidad. Si una respuesta puede producir una pérdida económica, afectar a un cliente, divulgar información privada o crear un compromiso importante, una persona debe verificarla y autorizar su uso antes de actuar.

Aplicación práctica: identifique qué podría salir mal antes de utilizar la respuesta

TEMA 2

Protección de datos e información del negocio

La información utilizada en una herramienta de IA debe limitarse a lo estrictamente necesario para la tarea. Antes de copiar un documento, una tabla o una conversación, conviene revisar si contiene datos personales, información financiera, contraseñas, secretos comerciales o datos que el negocio no está autorizado a compartir.

TIPO DE INFORMACIÓN	PREGUNTA DE CONTROL	EJEMPLO
Personal	¿Identifica a una persona?	Nombre, teléfono o documento de identidad.
Financiera	¿Expone dinero o medios de pago?	Cuenta bancaria o tarjeta.
Acceso	¿Permite entrar a un sistema?	Contraseña, código o token.
Clientes	¿Incluye datos privados de terceros?	Historial de compra o dirección.
Comercial	¿Es información interna sensible?	Costos, márgenes o estrategia no pública.
Laboral	¿Contiene información de empleados?	Evaluaciones o datos de nómina.
Pública	¿Ya está autorizada para divulgarse?	Horario, catálogo o información publicada.

Clasificar la información antes de compartirla

En lugar de introducir nombres, teléfonos o números de documentos, utilice etiquetas como Cliente 01 o Pedido A cuando la identidad no sea necesaria. Si una tarea puede resolverse con datos resumidos o anonimizados, esa opción reduce el riesgo y facilita un uso más responsable.

La empresa debe respetar sus políticas internas y la normativa aplicable sobre privacidad y protección de datos. Cuando exista duda sobre si una información puede utilizarse, lo prudente es no compartirla hasta recibir autorización o sustituirla por una versión anonimizada.

TEMA 3

Verificar resultados antes de utilizarlos

Una herramienta de inteligencia artificial puede producir respuestas convincentes que contengan errores, datos inexistentes o conclusiones que no se desprenden de la información suministrada. Por esta razón, la verificación debe formar parte del proceso normal de trabajo y no considerarse una actividad opcional.

Proceso en seis pasos

- 1 Identificar:** determine cuáles datos de la respuesta pueden comprobarse objetivamente.
- 2 Comparar:** contraste cifras, nombres, fechas y condiciones con la fuente original.
- 3 Separar:** distinga hechos proporcionados, inferencias y recomendaciones generadas.
- 4 Detectar faltantes:** revise si la IA asumió información que nunca fue suministrada.
- 5 Corregir:** elimine errores y solicite una nueva versión cuando sea necesario.
- 6 Autorizar:** utilice el resultado únicamente después de la revisión correspondiente.

Plantilla de verificación

Revisa esta respuesta utilizando únicamente los datos que te proporcioné. Separa en tres grupos: datos confirmados, información que debe verificarse y supuestos o interpretaciones. No agregues información nueva. Señala cualquier cifra, fecha, nombre o condición que no pueda confirmarse con los datos disponibles.

Señales de alerta

- Aparecen cifras, nombres, fuentes o fechas que usted no proporcionó ni puede comprobar.
- La respuesta expresa certeza aunque la información disponible sea incompleta.
- Se presentan recomendaciones importantes sin explicar los datos en que se apoyan.
- El resultado contradice documentos, registros o condiciones conocidas del negocio.

TEMA 4

Sesgos, transparencia y comunicación responsable

Las respuestas de una IA pueden reflejar información incompleta, patrones sesgados o criterios que no son apropiados para todas las personas. Esto es especialmente importante cuando la herramienta se utiliza para clasificar clientes, preparar mensajes, comparar personas o apoyar decisiones que puedan generar un trato desigual.

Prácticas responsables

SITUACIÓN	RIESGO	CONTROL
Clasificar clientes	Usar criterios injustificados o sensibles.	Utilizar variables pertinentes y revisión humana.
Redactar publicidad	Afirmar beneficios que no están comprobados.	Publicar solo información verificable.
Responder reclamos	Prometer soluciones no autorizadas.	Usar plantillas y aprobación cuando corresponda.
Analizar personas	Inferir características privadas.	No deducir atributos personales innecesarios.
Resumir opiniones	Confundir una opinión con un hecho.	Separar evidencia de interpretación.
Crear contenido	Presentar material generado como fuente experta.	Revisar, adaptar y aclarar límites cuando sea necesario.

Ejemplo con secciones

*OBJETIVO: resumir comentarios de clientes. DATOS: [COMENTARIOS ANONIMIZADOS].
FORMATO: tabla con tema, frecuencia y ejemplo resumido. CONTROLES: no infieras edad, género, origen, nivel económico ni otras características personales; separa lo que aparece en los datos de cualquier interpretación.*

También es importante ser transparente cuando un contenido generado con IA pueda confundirse con información humana, técnica o experta. La empresa debe evitar presentar como hecho una conclusión que solo es una sugerencia de la herramienta y debe aclarar límites cuando sean relevantes para el cliente.

Aplicación práctica: revise si su uso de IA puede producir un trato desigual o una impresión engañosa

TEMA 5

Supervisión humana y niveles de riesgo

No todas las tareas requieren el mismo nivel de control. Una idea para redes sociales tiene menos riesgo que una decisión sobre pagos, contratos o personal. Clasificar el impacto ayuda a definir cuánto control y revisión necesita cada uso de IA.

Ciclo de supervisión

- 1 **Clasificar:** determine si el impacto es bajo, medio o alto.
- 2 **Limitar:** defina qué puede hacer la IA y qué queda fuera de su alcance.
- 3 **Revisar:** asigne una persona responsable de comprobar el resultado.
- 4 **Escalar:** consulte a un responsable o profesional cuando sea necesario.
- 5 **Registrar:** documente decisiones, correcciones y versión utilizada.

Niveles orientativos

PREGUNTA	DECISIÓN DE CONTROL
¿Un error puede causar perjuicio?	Aumentar revisión y autorización.
¿Usa datos sensibles?	Minimizar, anonimizar o no compartir.
¿Afecta dinero, derechos o seguridad?	Mantener decisión humana.
¿Es difícil de verificar?	Buscar fuentes y revisión adicional.
¿Puede probarse sin impacto?	Realizar una prueba controlada.

Criterios de supervisión

- **Bajo:** borradores, ideas y organización de información no sensible; revisión habitual.
- **Medio:** mensajes a clientes o análisis internos; revisión obligatoria.
- **Alto:** decisiones financieras, legales, laborales o de seguridad; decisión humana y, cuando corresponda, revisión especializada.
- **Trazable:** debe poder identificarse qué datos e instrucciones originaron el resultado.

TEMA 6

Evaluar un resultado antes de aprobar su uso

La evaluación final reúne los controles del módulo. Antes de utilizar una respuesta, la empresa debe comprobar si es correcta, pertinente, segura, respetuosa de la privacidad y adecuada para la finalidad prevista. Una respuesta que no supera estos criterios debe corregirse o descartarse.

Ficha de evaluación

Tarea: [ACTIVIDAD]. Fuente de datos: [ORIGEN]. Resultado generado: [SALIDA]. Exactitud: [VERIFICADA / PENDIENTE]. Datos sensibles: [NO / SÍ Y CONTROL]. Riesgo de impacto: [BAJO / MEDIO / ALTO]. Responsable de revisión: [PERSONA]. Decisión: [APROBAR / CORREGIR / DESCARTAR].

Preguntas antes de aprobar

¿Los datos pueden verificarse? ¿La respuesta cumple exactamente la tarea? ¿Expone información que no debería compartirse? ¿Incluye suposiciones o afirmaciones no comprobadas? ¿Podría perjudicar a una persona o al negocio? ¿Existe un responsable que autorice el uso?

Documentar y aprender

Cuando se detecte un error, conviene registrar qué ocurrió y modificar la instrucción, los datos o el proceso para evitar que se repita. Si un uso deja de ser seguro, útil o verificable, debe suspenderse hasta que existan controles adecuados.

Regla de decisión

Aprobar cuando la información esté verificada y el riesgo sea aceptable; corregir cuando el problema pueda solucionarse antes de usarla; descartar cuando no sea posible verificar el contenido, proteger los datos o mantener una supervisión adecuada.

Gestión responsable

- Defina reglas sencillas sobre qué datos pueden utilizarse y cuáles están prohibidos.
- Conserve las instrucciones aprobadas y los criterios de revisión para tareas recurrentes.
- Capacite al equipo para reconocer errores, sesgos, datos sensibles y situaciones de alto riesgo.
- Revise periódicamente el uso de IA y ajuste los controles cuando cambien herramientas o procesos.

CIERRE

Evaluación responsable antes de utilizar IA

Ideas esenciales

- La inteligencia artificial puede equivocarse, toda información importante debe verificarse antes de utilizarse.
- Los datos personales, financieros, de acceso y comerciales sensibles deben protegerse y minimizarse.
- Los resultados deben revisarse para detectar información inventada, supuestos, sesgos y afirmaciones no comprobadas.
- El nivel de supervisión debe aumentar cuando la tarea tenga mayor impacto o utilice información más sensible.
- La responsabilidad final sobre las decisiones y compromisos del negocio permanece en las personas.

Actividad final

Seleccione una tarea real de su emprendimiento en la que utilice o podría utilizar IA. Identifique los datos necesarios, clasifique el nivel de riesgo y prepare una instrucción que incluya restricciones. Genere una respuesta de prueba y aplique la ficha de evaluación: verifique datos, detecte supuestos, revise privacidad, valore posibles sesgos y decida si el resultado debe aprobarse, corregirse o descartarse. Documente su decisión y explique qué control humano mantendrá.

Lista de comprobación

- **Datos mínimos:** solo se utiliza la información necesaria para la tarea.
- **Privacidad protegida:** no se incluyeron datos sensibles sin una razón y autorización adecuadas.
- **Exactitud verificada:** cifras, nombres, fechas, condiciones y afirmaciones importantes fueron comprobados.
- **Sesgos revisados:** se evaluó si el resultado puede producir un trato injusto o una conclusión engañosa.
- **Supervisión definida:** existe una persona responsable de revisar y autorizar el uso.
- **Decisión documentada:** el resultado fue aprobado, corregido o descartado con una razón clara.

Fuentes y recursos para ampliar

Recurso oficial: [NIST - Marco de Gestión de Riesgos de Inteligencia Artificial](#)

Recurso oficial: [OECD.AI - Principios de la OCDE sobre inteligencia artificial](#)

Recurso oficial: [OpenAI - Controles de datos y privacidad en ChatGPT](#)