

**Al para todos México**

**9:53**

**ESTAMOS POR COMENZAR...**



# IA ética y eficaz: Construyamos un marco de gobernanza en la empresa

**Samantha Villarreal**

**Microsoft MVP**



# CONTENIDO

- 01** Introducción a la gobernanza de la IA
- 02** Supervisión humana
- 03** Construcción de un marco de gobernanza
- 04** Preguntas

**TEMA**

# Introducción a la Gobernanza de la IA

# Gobernanza



**Sistema de Reglas**



**Prácticas**



**Procesos**

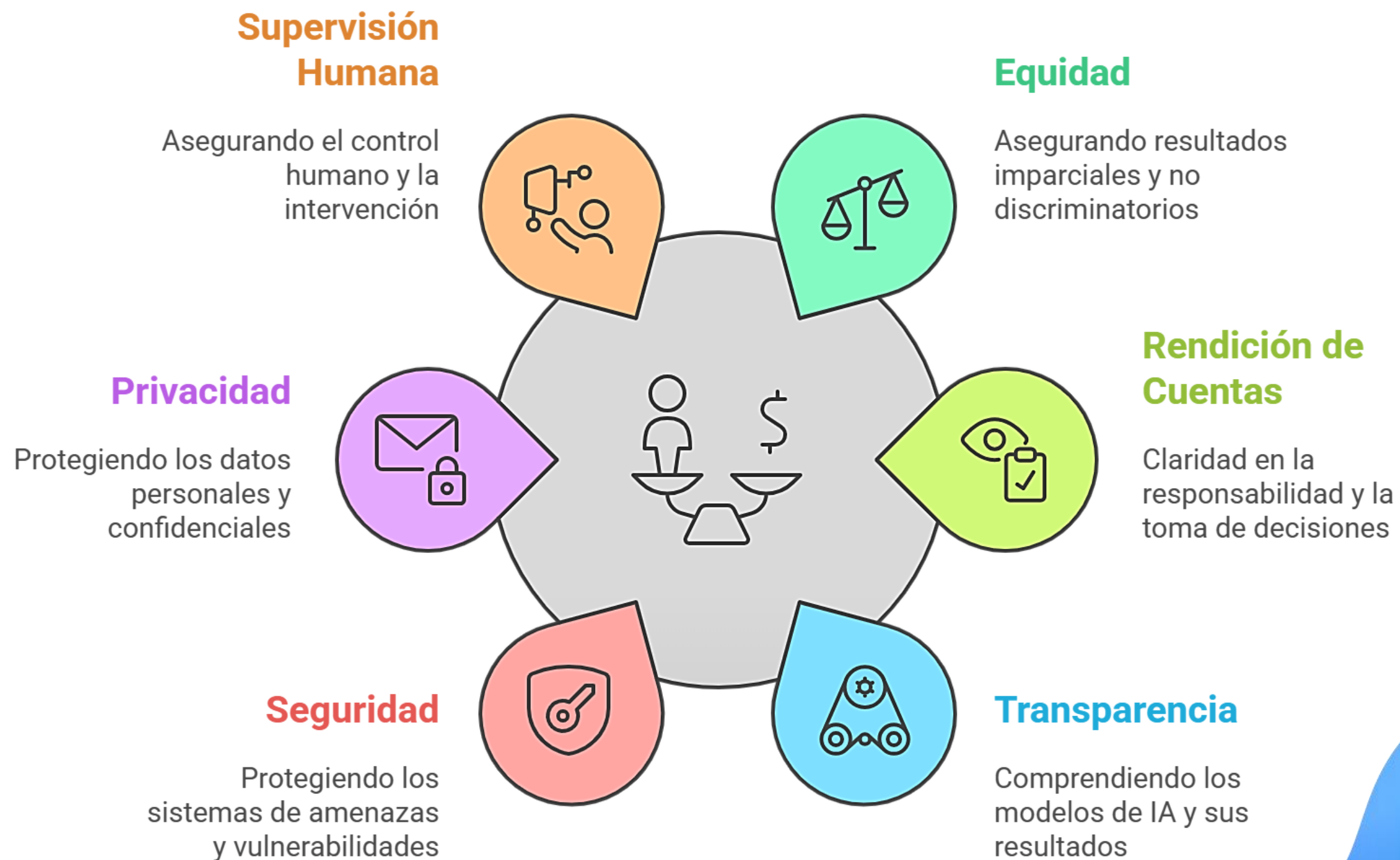


# Importancia

- **Capacidades Ampliadas**  
Creación de contenido, código, análisis, automatización.
- **Riesgos Inherentes**  
Desinformación, sesgos, privacidad de datos, propiedad intelectual, dependencia excesiva.
- **Necesidad Crítica**  
Establecer límites para maximizar beneficios y minimizar daños.



# Principios fundamentales



# Beneficios

- Mitigación de riesgos
- Fomento de confianza
- Impulso a la innovación
- Cumplimiento normativa
- Optimización del valor



**TEMA**

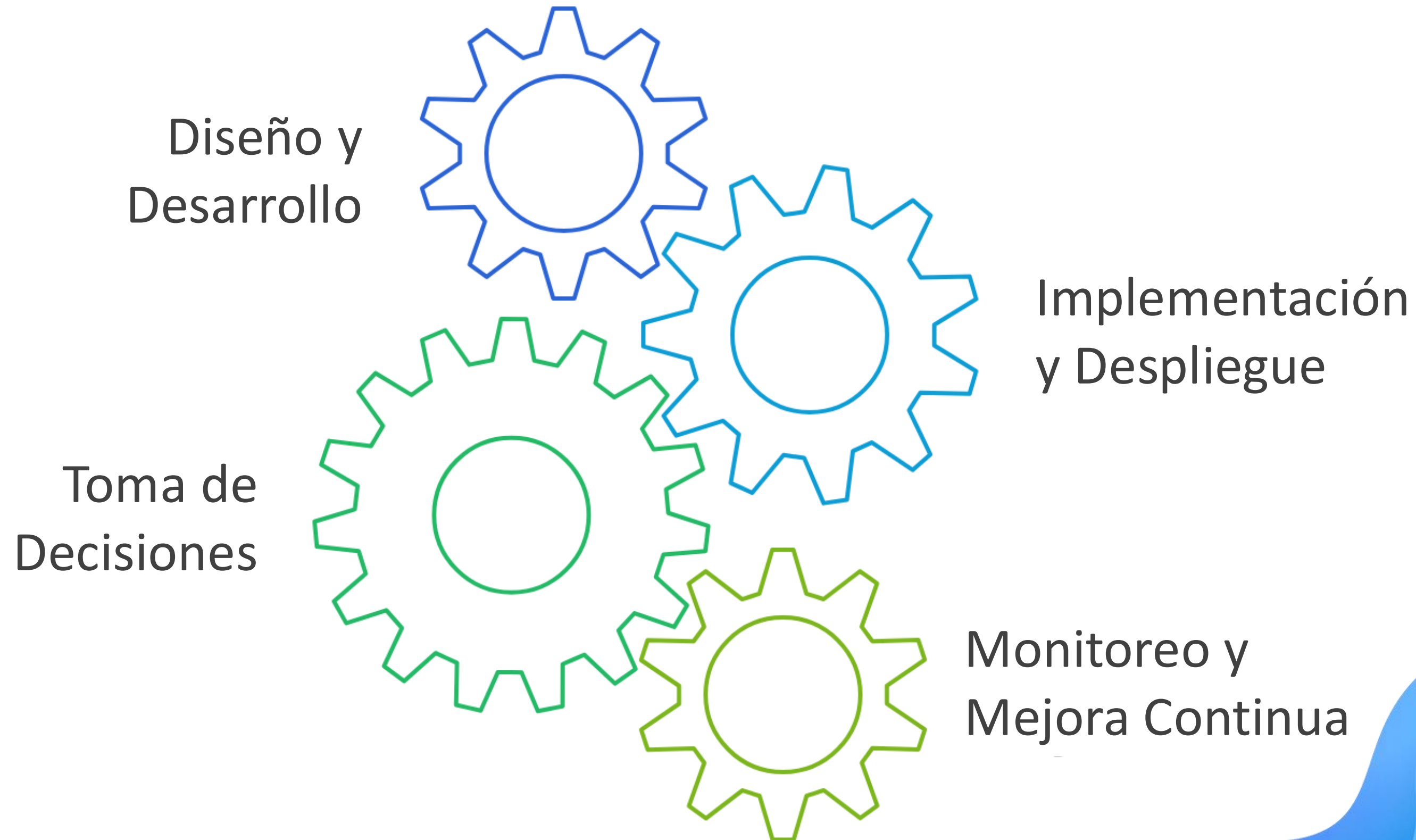
# Supervisión humana

# Rol de la supervisión humana

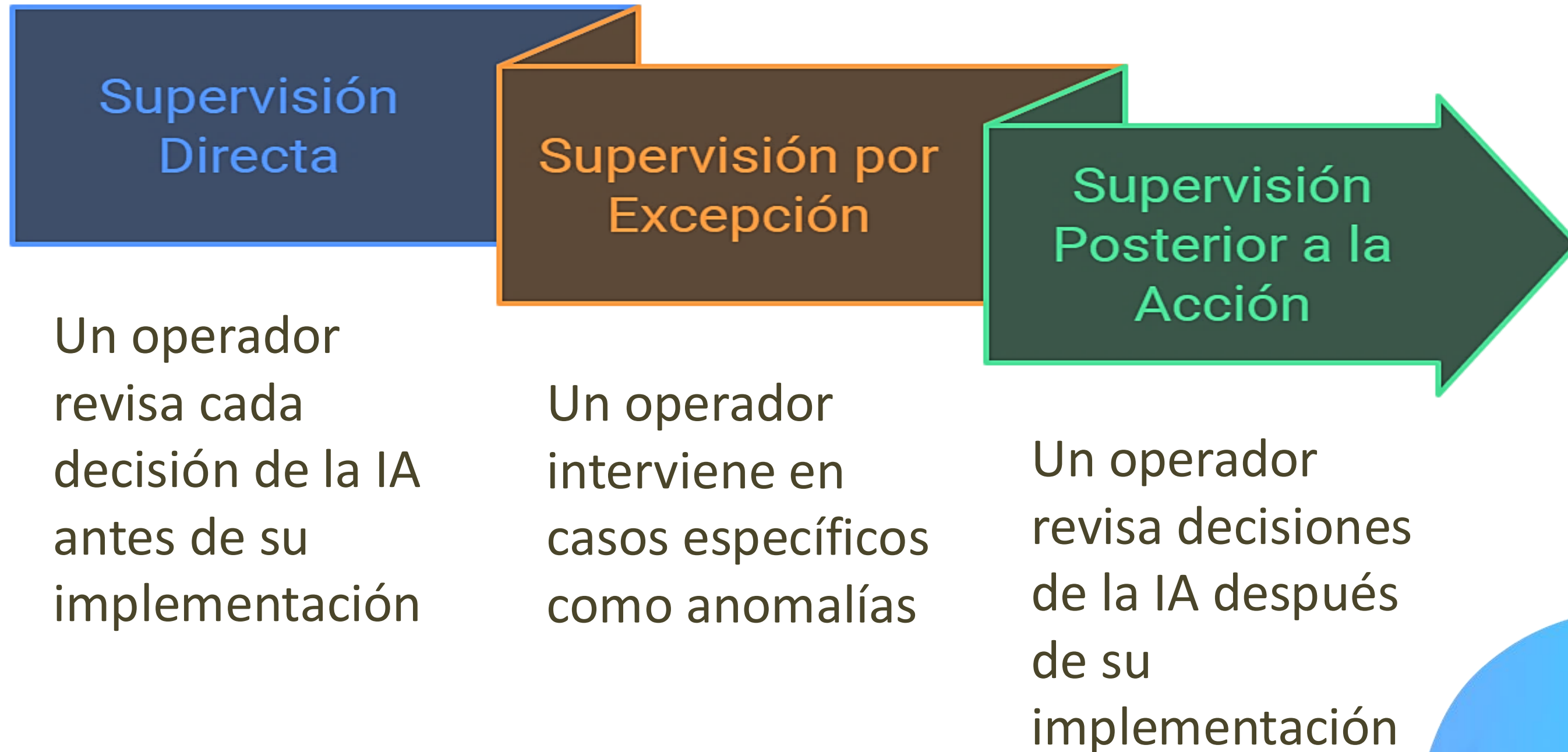
- Validar la precisión y relevancia de los resultados.
- Detectar y corregir sesgos.
- Asegurar la alineación con el contexto y los objetivos.
- Tomar decisiones finales en situaciones críticas o ambiguas.
- Garantizar la responsabilidad.



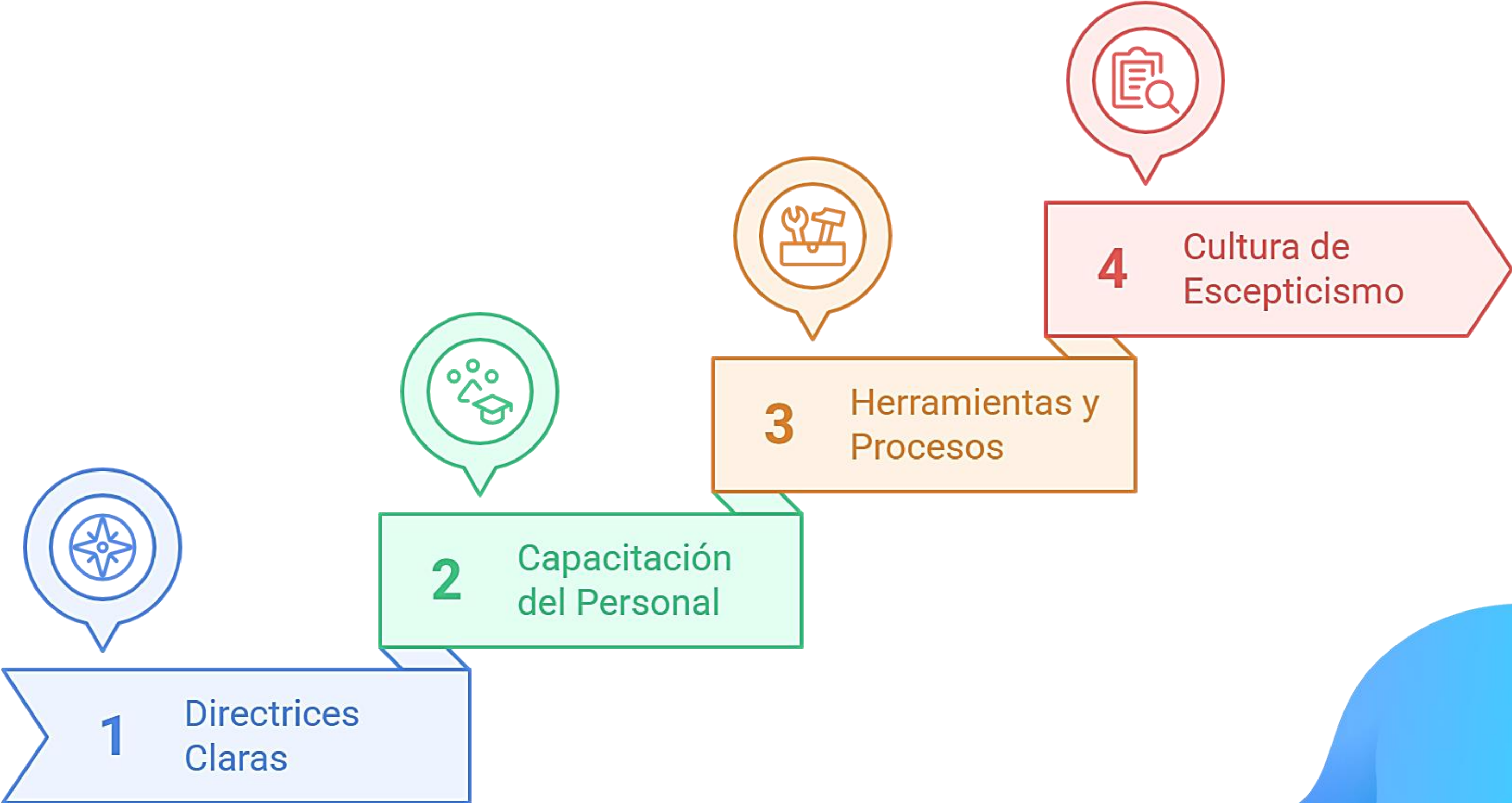
# Human in the loop



# Niveles de supervisión



# Implementación práctica



**TENDREMOS UN  
BREAK DE  
10 MINUTOS**

**YA VOLVEMOS...**

**9:53**

**TEMA**

# Construcción de un marco de gobernanza

# Necesidad de un marco formal

- Traduce principios y necesidad de supervisión en políticas y procedimientos.
- Proporciona estructura, claridad y consistencia en la organización.
- Es documentación que evoluciona con la tecnología y la experiencia.




# Políticas clave

-  **Uso Aceptable**  
Directrices sobre tareas permitidas y prohibidas.
-  **Datos y Privacidad**  
Normas sobre el manejo de datos sensibles.
-  **Seguridad**  
Requisitos para un uso seguro de la tecnología.
-  **Propiedad Intelectual**  
Reglas sobre autoría y plagio en contenido generado.
-  **Transparencia**  
Normas sobre la divulgación del uso de IA.



# Políticas clave




**Uso Aceptable**  
Directrices sobre tareas permitidas y prohibidas.




**Datos y Privacidad**  
Normas sobre el manejo de datos sensibles.




**Seguridad**  
Requisitos para un uso seguro de la tecnología.




**Propiedad Intelectual**  
Reglas sobre autoría y plagio en contenido generado.




**Transparencia**  
Normas sobre la divulgación del uso de IA.



**Tareas Permitidas**  
Actividades que se pueden realizar con la tecnología.



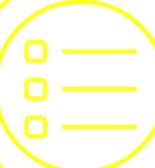
**Tareas Prohibidas**  
Actividades que no se deben realizar con la tecnología.




**Versiones Recomendadas**  
Opciones de versiones a utilizar según el contexto.



# Políticas clave




**Uso Aceptable**  
Directrices sobre tareas permitidas y prohibidas.




**Datos y Privacidad**  
Normas sobre el manejo de datos sensibles.




**Seguridad**  
Requisitos para un uso seguro de la tecnología.




**Propiedad Intelectual**  
Reglas sobre autoría y plagio en contenido generado.



**Transparencia**  
Normas sobre la divulgación del uso de IA.



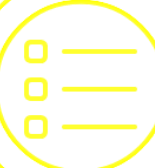
**Restricciones de Datos**  
Directrices sobre qué datos no se pueden introducir.



**Directrices de Anonimización**  
Instrucciones para anonimizar o generalizar la información.




# Políticas clave




**Uso Aceptable**  
Directrices sobre tareas permitidas y prohibidas.




**Datos y Privacidad**  
Normas sobre el manejo de datos sensibles.



**Seguridad**  
Requisitos para un uso seguro de la tecnología.




**Propiedad Intelectual**  
Reglas sobre autoría y plagio en contenido generado.



**Transparencia**  
Normas sobre la divulgación del uso de IA.



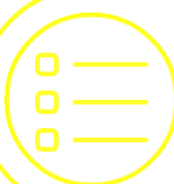
**Requisitos de Uso**  
Directrices para un uso seguro de tecnología.



**Gestión de Cuentas**  
Estrategias para manejar accesos y cuentas.




# Políticas clave




**Uso Aceptable**  
Directrices sobre tareas permitidas y prohibidas.




**Datos y Privacidad**  
Normas sobre el manejo de datos sensibles.




**Seguridad**  
Requisitos para un uso seguro de la tecnología.




**Propiedad Intelectual**  
Reglas sobre autoría y plagio en contenido generado.



**Transparencia**  
Normas sobre la divulgación del uso de IA.



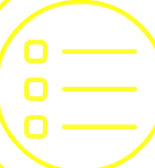
**Claridad sobre la autoría**  
Directrices claras sobre la propiedad del contenido.



**Obligación de Plagio**  
Requisito de revisar el contenido para prevenir el plagio.



# Políticas clave




**Uso Aceptable**  
Directrices sobre tareas permitidas y prohibidas.




**Datos y Privacidad**  
Normas sobre el manejo de datos sensibles.




**Seguridad**  
Requisitos para un uso seguro de la tecnología.



**Propiedad Intelectual**  
Reglas sobre autoría y plagio en contenido generado.



**Transparencia**  
Normas sobre la divulgación del uso de IA.



**Uso de IA**  
¿Cuándo se debe indicar que se utilizó IA para generar contenido o asistir en una tarea?



# Órgano de gobernanza

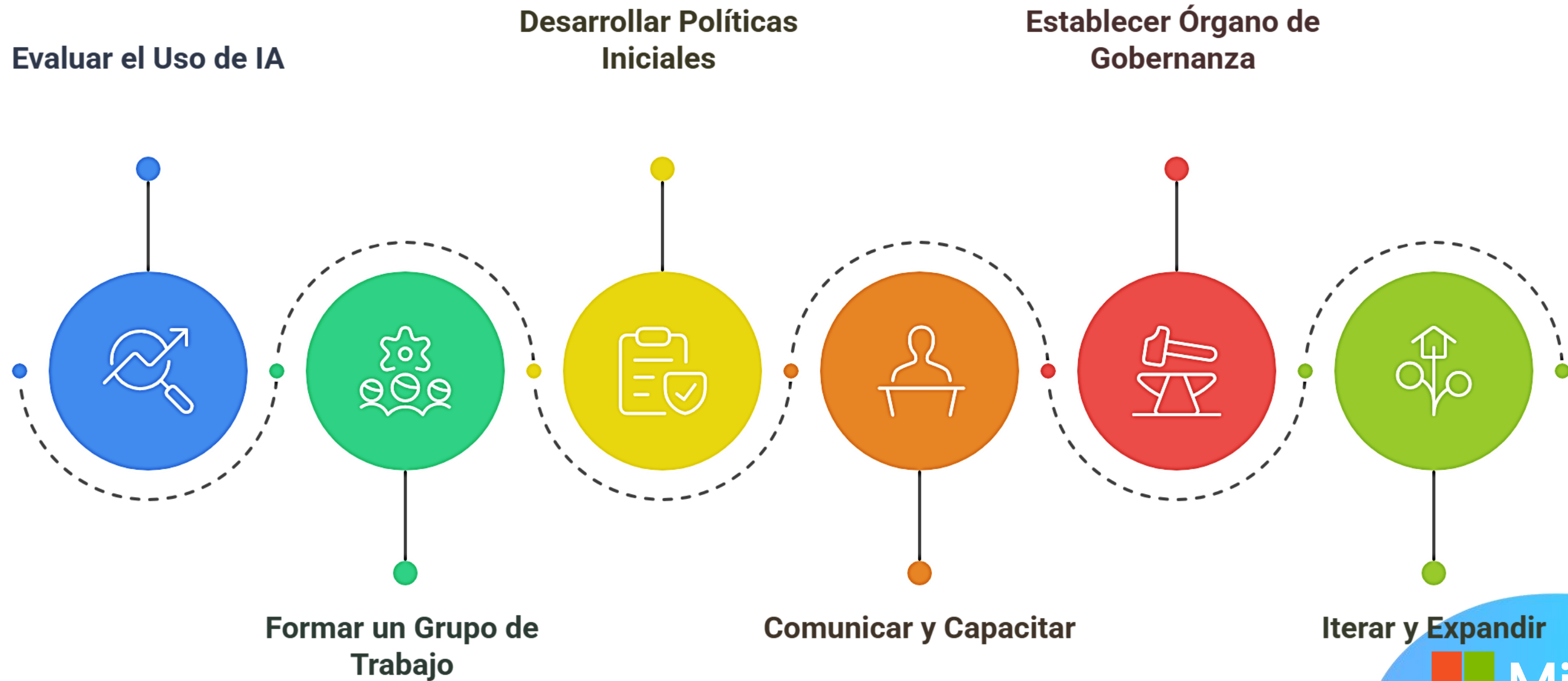
- Equipo multidisciplinario responsable de desarrollar, implementar y supervisar el marco de gobernanza.
- Puede ser un Consejo de Gobernanza de IA, o integrarse en estructuras existentes.
- Composición: Representantes de Legal, Cumplimiento, Riesgos, IT/Seguridad, Datos, Negocio, RRHH, y Ética.



# Consejo de Gobernanza



# Pasos concretos



# Sintetizando...

- La Gobernanza requiere un marco formal.
- Políticas Clave: Uso Aceptable, Datos/Privacidad, Seguridad, IP, Transparencia.
- Órgano de Gobernanza: Multidisciplinario, con roles y responsabilidades claras.
- Enfoque Práctico: Empezar por lo urgente, comunicar e iterar.



**TEMA**

**Preguntas**

# CÓDIGO DE EVALUACIÓN



¡MUCHAS GRACIAS!

