

## Cuestionario de autoevaluación de requisitos éticos de un sistema de IA

El cuestionario que se presenta a continuación pretende facilitar una guía básica para la autoevaluación por parte de los responsables de proyectos de IA<sup>1</sup>, del grado de cumplimiento de los requisitos éticos esenciales que cualquier sistema de IA debería satisfacer en línea con los principales pronunciamientos internacionales y, en particular, con los de la Comisión Europea (2019).

Las preguntas que aquí se plantean están orientadas a la evaluación de un sistema concreto<sup>2</sup>, y por ello, recomendamos a los responsables ofrecer las respuestas en coordinación con las áreas de la empresa o la institución que están a cargo de los datos, los modelos, las infraestructuras, las interfaces, etc. Así mismo, recomendamos mantener informada a la alta dirección para la resolución de posibles dilemas. Si el sistema de IA se contrata a un proveedor, muchas de las preguntas aquí planteadas deberían ser contestadas conjuntamente o por dicho proveedor en algunos casos.

En ningún caso hemos pretendido elaborar una relación exhaustiva de todas las preguntas que podrían formularse sino más bien una invitación amigable, de ahí su concisión, a la reflexión acerca del enfoque del sistema de IA objeto de análisis. Por esta razón las preguntas relacionadas con la seguridad y privacidad se han omitido intencionadamente al ser comunes a cualquier sistema. Respecto a la puntuación a asignar a cada respuesta, se sugiere: 0=no, 1=muy bajo .... 5=muy alto.

El cuestionario ha sido elaborado por el grupo de trabajo sobre ética y gobernanza del dato formado por los siguientes miembros del Observatorio: Enrique Bonsón (AECA), presidente, Montserrat Álvarez (Mapfre), Tamer Davut (PwC), Esther de la Torre (BBVA), José Fernando Esteban (ATOS), Carlos Fernández (Informa), Francisco Maregil (Banco de España), Montserrat Rubio (ICAC), Beatriz San Miguel (Fujitsu) y Enrique Turrillo (BDO). Por último, expresar nuestro agradecimiento a los participantes en los programas Máster de Auditoría del Instituto de Estudios Bursátiles (IEB) y de la Universidad de Granada por sus valiosas aportaciones de cara a la optimización de las preguntas.

---

<sup>1</sup> La inteligencia artificial (IA) se refiere a sistemas diseñados por humanos que, con un objetivo complejo, actúan en el mundo físico o digital al percibir su entorno, interpretar los datos estructurados o no estructurados recopilados, razonando sobre el conocimiento derivado de estos datos y decidir las mejores acciones para tomar (de acuerdo con parámetros predefinidos) para lograr el objetivo dado. Los sistemas de IA también pueden diseñarse para aprender a adaptar su comportamiento al analizar cómo el entorno se ve afectado por sus acciones anteriores. Como disciplina científica, la IA incluye varios enfoques y técnicas, como el aprendizaje automático (de los cuales el aprendizaje profundo y el aprendizaje de refuerzo son ejemplos específicos), el razonamiento automático (que incluye planificación, programación, representación y razonamiento del conocimiento, búsqueda y optimización) y la robótica. (que incluye control, percepción, sensores y accionadores, así como la integración de todas las demás técnicas en sistemas ciber-físicos). *Comisión Europea, High Level Expert Group on AI (2019): «Una definición de la inteligencia artificial: Principales capacidades y disciplinas científicas».*

<sup>2</sup> Un sistema de inteligencia artificial está compuesto por los datos de origen, los modelos, las infraestructuras, las interfaces de usuario .... y el caso de uso para el que se pone todo ello en marcha.

| Cuestionario de autoevaluación de requisitos éticos de un sistema de IA   | Ptos. |
|---|-------|
| <b>Puntuación total de la autoevaluación</b>  |       |
| <b>1. Acción y supervisión humanas</b>  |       |
| En caso de que el sistema de IA tenga impacto en personas ¿Se ha realizado un análisis del impacto del sistema en sus derechos fundamentales?   |       |
| ¿Ha analizado si los objetivos y desarrollo de su Sistema de IA están alineados con las <b>pautas y/o valores éticos</b> de su compañía o institución?  |       |
| ¿Se ha establecido algún mecanismo para identificar los <b>posibles conflictos entre principios, intereses y/o valores éticos</b> que implica el sistema de IA y evaluarlos?  |       |
| ¿Ha tenido usted en cuenta si la <b>asignación de tareas entre el sistema de IA y los empleados</b> garantiza la mejora de las capacidades humanas y que las interacciones sean satisfactorias?                         |       |
| ¿Ha establecido mecanismos para determinar el grado adecuado de <b>intervención humana en función del caso de uso</b> ?   |       |
| <b>2. Transparencia</b>   |       |
| ¿Ha establecido el nivel de <b>interpretabilidad/explicabilidad</b> que exige el caso de uso del sistema?   |       |
| <b>Si el caso de uso lo requiere</b> ¿Ha evaluado en qué medida puede explicarse la lógica de los resultados y <b>cómo elaborar esta explicación</b> ?  |       |
| ¿Se han <b>documentado</b> las fuentes de datos y los métodos usados para <b>diseñar, probar y desarrollar</b> el sistema algorítmico, y sus <b>resultados</b> ?  |       |
| ¿Ha <b>facilitado a los usuarios finales</b> conocer que están <b>interactuando con un sistema de IA</b> y no con otro ser humano?  |       |
| <b>3. Diversidad, no discriminación y equidad</b>   |       |
| ¿Existe una <b>estrategia (procedimiento)</b> para <b>evitar</b> crear o reforzar <b>un sesgo injusto</b> en el sistema de IA, tanto en relación con los datos de entrada como en lo referente al diseño del algoritmo? |       |
| ¿Ha puesto en marcha <b>medidas para evitar sesgos injustos en los resultados</b> estableciendo las medidas adecuadas de control y supervisión?   |       |
| ¿Ha estudiado la posibilidad de introducir algún mecanismo para <b>incorporar la participación de diferentes perfiles de personas</b> en el desarrollo y la utilización del sistema? Ej. Test de usuarios               |       |
| <b>4. Bienestar social y ambiental</b>  |       |
| ¿Se está usando la <b>IA en su compañía, también para el bien social?</b> (Por ejemplo, si existe algún proyecto enfocado a cumplir los objetivos de desarrollo sostenible)   |       |
| ¿Ha evaluado si el <b>impacto social y ambiental</b> del desarrollo de la IA en su compañía o institución está <b>alineado con los Objetivos de Desarrollo Sostenible</b> de las Naciones Unidas?                       |       |
| <b>5. Rendición de cuentas</b>  |       |
| ¿Existen mecanismos para facilitar la <b>auditabilidad del sistema</b> por parte de agentes internos o independientes?  |       |
| ¿Existen <b>responsabilidades</b> individuales claramente <b>identificadas</b> en cada fase?  |       |
| ¿Ha establecido un conjunto de <b>mecanismos adecuado que permita obtener compensación a terceros</b> en el caso de que se produzca cualquier daño o efecto adverso?  |       |