

HERRAMIENTAS DE ALGORITMOS ÉTICOS

- El **GobLab UAI**, laboratorio de innovación pública de la Escuela de Gobierno de la Universidad Adolfo Ibáñez, ha trabajado desde 2020 en promover el uso responsable y transparente de la inteligencia artificial y sistemas automatizados o semiautomatizados en el sector público chileno, buscando con ello apoyar a las instituciones públicas innovadoras.

En la plataforma Algoritmos Públicos se encuentran disponibles tres herramientas desarrolladas por el GobLab UAI. Estas son gratuitas, estandarizadas, de código abierto y en español. Permiten evaluar, prevenir y mitigar riesgos éticos en los modelos de ciencia datos e inteligencia artificial, es decir, **apoyan la incorporación de estándares éticos y responsables durante todo el ciclo de vida de un proyecto de IA y ciencia de datos.**

Socios:



Patrocina:

Financia:

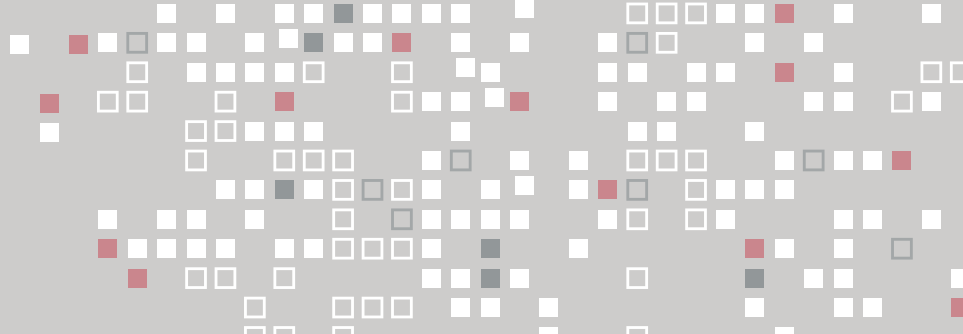


Ciclo de vida de un proyecto de IA y ciencia de datos

Etapas sugeridas para el uso de cada herramienta



Encuentra las herramientas de algoritmos éticos en
www.algoritmospublicos.cl



Herramientas

Evaluación de Impacto Algorítmico

Identifica riesgos éticos desde una mirada técnica, legal y de gestión, y entrega recomendaciones de mitigación personalizadas al proyecto de IA.

1

Medición de Sesgos y Equidad

Mide e identifica sesgos en las predicciones del modelo de IA, para crear sistemas más justos y prevenir la discriminación.

2

Ficha de Transparencia del Modelo

Documenta y comunica de manera clara la lógica de decisión detrás del sistema de IA y explica aspectos claves para promover la transparencia.

3

¿Para qué sirven estas herramientas?

- Para una implementación ética, responsable y segura de proyectos de IA y ciencia de datos.
- Para dar cumplimiento a los objetivos de la Política Nacional de Inteligencia Artificial.
- Para aplicar la circular Lineamientos para el uso de herramientas de inteligencia artificial en el sector público, del Ministerio de Ciencia y la Secretaría de Gobierno Digital.
- Para apoyar la implementación de las recomendaciones de transparencia algorítmica en servicios públicos del Consejo para la Transparencia.

¿Quiénes las pueden utilizar?

- Instituciones públicas, sector privado, ONGs, empresas de tecnología.

1 Evaluación de impacto algorítmico

¿Por qué utilizar esta herramienta?

- Permite anticipar riesgos y dimensionar el grado de impacto de un sistema de IA.
- Ayuda a identificar a los **responsables** de la implementación y operación de sistemas de IA, asegurando un uso ético y transparente.
- Fomenta la **confianza pública** y **previene riesgos** como el sesgo o la discriminación, infracciones de protección de datos y opacidad.

¿En qué consiste la herramienta?

- Realiza un análisis detallado para identificar riesgos éticos, sociales, técnicos y sus posibles daños, permitiendo tomar medidas concretas para evitarlos o mitigarlos.
- Recomienda pasos a seguir y herramientas para tomar acciones que permitan mitigar los riesgos éticos del sistema.

¿Cuál es el resultado de la herramienta?

Informe con nivel de impacto del sistema de IA y recomendaciones personalizadas para mitigar riesgos en:

- Proporcionalidad
- Normativa
- Protección de datos
- Licencia social
- Gobernanza
- Rendición de cuentas
- Transparencia
- Ciberseguridad
- No discriminación - equidad

2 Medición de sesgos y equidad

¿Por qué utilizar esta herramienta?

- Para implementar modelos de IA que sean más justos y equitativos con las personas.
- Para cumplir con la legislación chilena actual. La Ley N° 20.609, establece la no discriminación arbitraria para variables protegidas como sexo, nacionalidad, situación socioeconómica, entre otras.
- Porque permite incorporar la dimensión de equidad en la evaluación del desempeño de un modelo de IA, considerando los errores que tiene en diferentes grupos de la población.

¿En qué consiste la herramienta?

- Mide los sesgos en las predicciones de los modelos de clasificación binaria, permitiendo entender las diferencias que existen en el desempeño de un sistema.
- Permite ajustar la medición de sesgos a la realidad del sistema de IA, escogiendo las variables protegidas a evaluar y el porcentaje de sesgo que sea aceptable para el proyecto. Las variables protegidas son características de las personas que están legal o éticamente resguardadas contra la discriminación arbitraria.
- Se puede utilizar durante la etapa de evaluación del modelo o para auditar los sesgos de sistemas existentes.

¿Cuál es el resultado de la herramienta?

Reporte con los resultados de la medición de sesgos en las variables protegidas del modelo de IA, usando nueve métricas disponibles:

- Tasa de verdaderos positivos
- Falsos negativos/Tamaño del grupo
- Tasa de falsa omisión
- Tasa de falsos negativos
- Paridad de selección igualitaria
- Paridad demográfica
- Falsos positivos/Tamaño del grupo
- Tasa de falsos descubrimientos
- Tasa de falsos positivos

3

Ficha de transparencia algorítmica

¿Por qué utilizar esta herramienta?

- Transparenta información clave sobre un modelo de ciencia de datos o IA, lo que fomenta la confianza, la rendición de cuentas y la responsabilidad en la toma de decisiones.
- Ayuda a documentar el modelo de ciencia de datos o IA de manera estructurada.
- Permite a las personas acceder a información valiosa sobre cómo la IA participa en la toma de decisiones que las afectan, lo que mejora la calidad de la democracia.

- Permite cumplir con las Recomendaciones de Transparencia Algorítmica del Consejo para la Transparencia.

¿En qué consiste la herramienta?

- Resume la arquitectura del modelo, los datos de entrenamiento, las métricas de rendimiento, **posibles riesgos éticos** y otros aspectos cruciales para comprender y evaluar su funcionamiento.
- Aborda preguntas de los distintos elementos, características y consideraciones del modelo.

¿Cuál es el resultado de la herramienta?

Documento claro y amigable con el usuario que transparenta:

- Detalles del modelo
- Visión general del modelo
- Clasificación
- Métricas de rendimiento
- Datos de entrenamiento
- Datos de evaluación
- Consideraciones éticas
- Advertencias y recomendaciones
- Reclamación



HERRAMIENTAS ALGORITMOS ÉTICOS

Encuentra las herramientas de algoritmos éticos en:

<https://algoritmospublicos.cl/herramientas>

Si quieres contactarte con nosotros escribe a goblab@uai.cl

Socios:



Patrocina:

Financia:

