

Propuesta de programa para Ética de la Inteligencia Artificial. (ONLINE)

Miguel Ramón Viguri Axpe.

Curso 2024/25

CAMPUS VIRTUAL INTERNACIONAL

Código Asignatura? – ÉTICA DE LA INTELIGENCIA ARTIFICIAL (ONLINE)

Información de la Guía Docente

Titulación:

¿?

Curso académico:

2024/25

Asignatura:

Código? - Ética de la Inteligencia Artificial

Periodo impartición:

Segundo semestre

Curso:

(los grupos que estén inscritos?)

Tipo:

Obigatoria/Optativa?

Créditos ECTS:

6.0

Modalidad:

ONLINE

JUSTIFICACIÓN

La velocidad con la que nuestra sociedad se digitaliza e incorpora en sus procesos de articulación aplicaciones, servicios y sistemas digitales, hace que la atención a estas dinámicas y sus efectos adquieran un interés creciente. Por norma general, los avances se han venido desarrollando impulsados por las oportunidades que ofrecen los mercados, atendiendo especialmente a motivaciones de naturaleza económica para la implantación de productos o servicios digitales. Especialmente significativo es el crecimiento en la capacidad que tienen estos sistemas, impulsados con sistemas de inteligencia artificial con funcionalidades cada vez

más avanzadas. En esta dinámica de crecimiento, surgen asimismo con fuerza nuevos cuestionamientos de naturaleza ética en relación con los efectos generados por este proceso de digitalización cada vez más rápido, de alcance global y con mayores impactos.

La ética se entiende como un saber práctico que, aunque sustentada en una base teórica, busca realizarse y aplicarse en todas las dimensiones de la vida moral concreta de la persona tanto en el ámbito de la ciudadanía, como en sus responsabilidades públicas a través del ejercicio profesional. Por ello, la asignatura de Ética de la IA pretende impartir una formación ética básica que permita a los ciudadanos/as usuarios/as de estas tecnologías identificar los factores fundamentales que garantizan que el uso de dichos productos cumple con unos estándares éticos mínimos y, de esta forma, proteger sus derechos cívicos y, sobre todo, los Derechos Humanos. Esta asignatura, asimismo, tendrá en cuenta especialmente la orientación profesional de quienes intervienen en la creación de productos basados en IA.

RESULTADOS DE APRENDIZAJE E INDICADORES DE EVALUACIÓN:

1. Resultados de Aprendizaje Transversales o Genéricos

R.A.T..1: Autoconocimiento, ética, responsabilidad social y medioambiental y apertura a la trascendencia

Actuar de manera ética, igualitaria, inclusiva, responsable y sostenible, con uno mismo, con los otros (hombres y mujeres para los demás), con la sociedad (justicia social) y con el planeta en su conjunto (medio ambiente), planteándose las grandes preguntas de la vida.

Indicadores:

- **R.A.T.1.1:** Aplicar modelos de ética deliberativa y prudencial para la solución de casos reales o simulados.

R.A.T..2: Aprendizaje profundo, autónomo y crítico

Actualizar los propios aprendizajes, cuestionando las formas habituales de actuación e interrogándose sobre las cuestiones con una perspectiva crítica, reflexionando sobre el propio conocimiento y la propia manera de aprender.

Indicadores:

- **R.A.T.2.1:** Cuestiona la realidad que le rodea analizando los principios, hechos, actitudes y valores implicados en situaciones y/o casos prácticos relacionados con tecnologías de IA.
- **R.A.T.2.2:** Realiza producciones personales a partir de materiales de aprendizaje trabajados que demuestren la propia forma de razonar y actuar.

0. Resultados de Aprendizaje Específicos

R.A.E.1: Comprensión de cómo aplicar enfoques éticos al análisis de impactos en el ámbito digital

Identificar y analizar éticamente los impactos generados por la implementación de las tecnologías digitales en distintos espacios sociales.

Indicadores:

- **R.A.E.1.1:** Analiza impactos de las tecnologías digitales y la IA en los ámbitos político y jurídico, identificando y mostrando las implicaciones éticas de dichos impactos.
- **R.A.E.1.2:** Analiza impactos de las tecnologías digitales y la IA en la esfera psicológica y en la construcción de la identidad individual y social, identificando y mostrando las implicaciones éticas de dichos impactos.

R.A.E.2: Conocimiento y análisis ético de técnicas y herramientas digitales vinculadas a la IA

Conoce y analiza éticamente técnicas y herramientas digitales vinculadas a la IA, con el fin de entender las características de la información registrada y sus posibilidades y riesgos de uso.

Indicadores:

- **R.A.E.2.1:** Analiza con perspectiva ética técnicas de procesamiento de datos digitales, algoritmia e Inteligencia Artificial, con el fin de determinar las tipologías de sus procesos y respuestas y sus posibilidades y riesgos de uso.
- **R.A.E.2.2:** Concibe estrategias y procedimientos de desarrollo tecnológico digital apropiados para una organización, producto o proceso, identificando los diferentes aspectos éticos implicados.

PRERREQUISITOS

No existen requisitos previos para cursar esta asignatura.

CONTENIDOS

Unidad 1.- Conceptos básicos y fundamentos de análisis ético:

- ¿Qué es la ética? Diferencia entre ética y moral
- Diferentes éticas
- Conceptos de justicia, responsabilidad, cuidado
- La ética como racionalidad práctica:
- La ética como instancia crítica.
- La reflexión ética como diálogo interdisciplinar y deliberación.
- Las tecnologías y la ciencia: saberes no-neutrales éticamente.
- Conceptos básicos de Big Data e IA.
- Introducción al impacto social y cultural de la Digitalización y la IA.

- Análisis ético de las tecnologías y sus productos.

Unidad 2.- La relación del individuo y la tecnología en la construcción social: la IA y su impacto personal y psicológico

- La toma de decisiones irracional
- Capacidades limitadas, heurísticos y sesgos
- Heurístico de la máquina y sesgo de automatización
- Sesgos en el ciclo de vida de la IA
- La automatización de decisiones
- Human-in-the-loop
- La mediación de la tecnología en las decisiones

Unidad 3.- Problemas y horizontes sociales: Impactos Políticos y Jurídicos de la IA

- Problemas referidos a la Libertad: Amenazas a la autorrealización y la emancipación mediante el uso de la IA. Libertad como participación, IA en las elecciones y libertad de expresión.
- Problemas referidos a la Igualdad y justicia: Igualdad de oportunidades y de derechos y autonomía (filosofía liberal anglosajona); Fragmentación y erosión de la cohesión comunitaria, Marginalización de culturas y prácticas locales (filosofía comunitarista). Problemas referidos a la Democracia.
- Principios del Derecho Constitucional y su sistema de garantías jurídicas.
- Desafíos de la IA al Derecho Civil y Administrativo. ¿Qué modificaciones son necesarias?

ESTRATEGIA DE ENSEÑANZA-APRENDIZAJE, SISTEMA DE EVALUACIÓN Y DOCUMENTACIÓN

Miguel Ramón Viguri Axpe	Castellano
ESTRATEGIA DE ENSEÑANZA-APRENDIZAJE	
<p>La asignatura se desarrollará, en formato virtual, siguiendo el Modelo Deusto para la Formación Online que tiene su origen en el Modelo de Aprendizaje de la Universidad Online de Deusto (MAUD) y ha sido desarrollado para dar respuesta a las necesidades de un estudiante en un contexto online.</p>	

La asignatura está organizada en 3 unidades. En todas ellas se parte de un escenario de aprendizaje que permitirá ir desarrollando las competencias definidas para la asignatura permitiendo la aplicación de los contenidos teóricos en la práctica.

De acuerdo a los 6 ECTS asignados, la dedicación requerida al estudiante para el seguimiento de la asignatura y el cumplimiento de sus requisitos es de 150 horas. La dedicación a la asignatura deberá ser de unas 10 horas semanales. Estas horas estarán distribuidas del siguiente modo:

- Resolución de retos planteados en los escenarios de aprendizaje: en cada unidad los estudiantes tendrán que resolver un escenario de aprendizaje, que será el que guíe todo su trabajo a lo largo de la unidad. 67 horas
- Lectura y estudio de materiales: el estudiante contará con diferentes recursos (lecturas, vídeos, actividades de aplicación inmediata) en cada una de las unidades que tendrá que revisar e interiorizar para completar la unidad. 80 horas.
- Tutorías individualizadas y de grupos. 3 horas

Uso de herramientas de IA en la asignatura

El uso de herramientas de inteligencia artificial (IA), o contenidos generados a través de las mismas, estará autorizado como recurso auxiliar en el desarrollo de los trabajos solicitados al estudiante en esta asignatura. En el caso de hacer uso de la IA obligatorio el hacer constar las fuentes correspondientes, y/o las herramientas digitales utilizadas y la descripción de su uso en la elaboración.

En concreto el estudiante debe mostrar:

- **Honestidad:** Debe siempre citarse el tipo de uso realizado de cada herramienta.
- **Autoría:** El autor humano debe seguir asumiendo la autoría del texto, reelaborando y completando todo lo que sea preciso.
- **Comprensión:** Debe comprenderse y poder defender la totalidad de las opiniones e informaciones finalmente redactadas.
- **Evaluación:** Debe aceptarse que las actividades de evaluación podrán incorporar comprobaciones de que el alumnado ha entendido adecuadamente la información obtenida de fuentes basadas en IA, y es capaz de defender y enriquecer esa información.
- **Verificación:**
 - Debe validarse la corrección/autenticidad de los datos suministrados por IA.

Miguel Ramón Viguri Axpe	Castellano
--------------------------	------------

- Deben utilizarse y referenciarse fuentes tradicionales, fiables y confirmadas.
- **Responsabilidad:** Deben entenderse los riesgos de suministrar información a las IAs.

Se recuerda que no está permitido copiar y/o plagiar, ni generar mediante una Inteligencia Artificial ningún trabajo académico sin tener como referencia los puntos señalados anteriormente. La comisión de este tipo de prácticas fraudulentas se considerará falta grave, en el sentido del art. 75 del Reglamento de Estudiantes (Boletín Oficial de la Universidad de Deusto nº 81, junio de 2023).

SISTEMA DE EVALUACIÓN

Se evaluarán los siguientes resultados de aprendizaje genéricos y específicos en la asignatura mediante las siguientes técnicas y procedimientos de calificación:

Competencias	Actividades y pruebas
R.A.G1: 10%	Unidad 1: 5% Unidad 2: 2.5% Unidad 3: 2.5%
R.A.G.2 10%	Unidad 1:5 % Unidad 2: 2.5% Unidad 3: 2.5%
R.A.E.1: 40%	Unidad 1: 20% Unidad 2: 10% Unidad 3: 10%
R.A.E.2: 40%	Unidad 1: 20% Unidad 2: 10% Unidad 3: 10%

Tomando como referencia las actividades de evaluación el peso final de cada una de ellas en la calificación final es el siguiente:

Actividad	%
Escenario de Aprendizaje 1	15%
Escenario de Aprendizaje 2	15%

Miguel Ramón Viguri Axpe	Castellano
---------------------------------	-------------------

Escenario de aprendizaje 3	20%
Escenario de aprendizaje 4	12.5%
Escenario de aprendizaje 5	12.5%
Escenario de aprendizaje 6	25%

CONVOCATORIA ORDINARIA

Para aprobar la asignatura será necesario aprobar cada uno de los escenarios de aprendizaje. Es decir, obtener al menos un 5 sobre 10 en cada uno de los escenarios.

CONVOCATORIA EXTRAORDINARIA

Para aprobar la asignatura en convocatoria extraordinaria tienen que cumplir dos requisitos:

- Aquellas personas que no hayan superado o entregado los escenarios de aprendizaje a lo largo de la evaluación continua deberán enviarlos en la fecha indicada, para que sean corregidos.
- El/la estudiante tiene que obtener un 5 sobre 10 en cada uno de los escenarios.

DOCUMENTACIÓN

DOCUMENTACIÓN BÁSICA

- Etxeberria, X. (2002). *Temas Básicos de Ética*. Desclée de Brouwer, Bilbao.
- Bilbao, G., Fuertes, J. y Guibert, J.M. (2006). *Ética para ingenieros*. Desclée de Brouwer, Bilbao.
- Villas Olmeda, M. y Camacho Ibáñez, J. (2022). *Manual de Ética Aplicada en Inteligencia Artificial*. Anaya, Madrid.

DOCUMENTACIÓN COMPLEMENTARIA

- Casado, E. G. (2021). El enfoque europeo de Inteligencia Artificial. *Revista de Derecho Administrativo*, (20), 268-289.
- Casado, Eduardo Gamero. "El enfoque europeo de Inteligencia Artificial." *Revista de Derecho Administrativo* 20 (2021): 268-289.

Miguel Ramón Viguri Axpe	Castellano
<ul style="list-style-type: none"> ● Arandia, P., Ley, M., Siqués, S., Martín, L., Ortega, M., Mateo, M., & Luengo, J. (24 de Abril, 2024). Riscanvi. Un algoritmo define el futuro de los presos en Cataluña: ahora sabemos cómo funciona. El Confidencial. https://www.elconfidencial.com/tecnologia/2024-04-24/riscanvi-algoritmo-cataluna-prisiones-presos-inteligencia-artificial_3871170/ ● Armentia, I. (16 de Noviembre, 2018). El algoritmo de Netflix nunca será el videoclub de tu barrio. eldiario.es. https://www.eldiario.es/euskadi/blogs/al-margen/algoritmo-netflix-videoclub-barrio_132_1835335.html ● Bellio, N. (5 de Abril, 2023). Los sesgos con los extranjeros en el algoritmo de violencia de género de la Ertzaintza. El Salto. https://www.elsaltodiario.com/policia/sesgos-extranjeros-del-algoritmo-violencia-genero-ertzaintza ● Belmonte, E. (16 de Mayo, 2019). La aplicación del bono social del Gobierno niega la ayuda a personas que tienen derecho a ella. Civio. https://civio.es/tu-derecho-a-saber/2019/05/16/la-aplicacion-del-bono-social-del-gobierno-niega-la-ayuda-a-personas-que-tienen-derecho-a-ella/ ● Digital Future Society. (2022). Hacia una supervisión significativa de los sistemas automatizados de toma de decisiones. https://digitalfuturesociety.com/app/uploads/2022/11/Hacia_una_supervision_significativa_de_los_sistemas_automatizados_de_toma_de_decisiones.pdf ● Hao, K. (11 de Noviembre, 2021). Caso práctico: probamos por qué un algoritmo judicial justo es imposible. MIT. https://www.technologyreview.es/s/13800/caso-practico-probamos-por-que-un-algoritmo-judicial-justo-es-imposible ● Pastor, J. (12 de Marzo, 2024). Internet tal y como la conocíamos está muriendo: la IA generativa se la está cargando. Xataka. https://www.xataka.com/robotica-e-ia/internet-tal-como-conociamos-esta-muriendo-ia-generativa-se-esta-cargando ● Saura, G., & Aragón, L. (2021). Un algoritmo impreciso condiciona la libertad de los presos. La Vanguardia. https://www.lavanguardia.com/vida/20211206/7888727/algoritmo-sirve-denegar-permisos-presos-pese-fallos.html 	