



PROYECTO ARTIN FUTURE

DETALLES DEL CURSO

Desarrollado por: 

BEST Institut für berufsbezogene Weiterbildung und
Personaltraining GmbH

ÍNDICE

Índice

A. Introducción	2
Sobre esta herramienta.....	2
Resumen.....	2
B. Información de la herramienta	2
C. Información sobre cada tema	5
1. ¿Qué es la IA? (IA débil, IA fuerte, superinteligencia artificial SIA)	6
1.1 Introducción a la inteligencia artificial.....	6
1.2 La IA más detalladamente	6
1.3 Principios éticos y su relevancia para las TIC e IA.....	7
2. ¿Qué hace la IA? El potencial y los riesgos de la IA desde una perspectiva ética.....	8
2.1 El potencial de la IA	8
2.2 Los riesgos de la IA	9
2.3 Estudio de caso del potencial y riesgos de la IA	9
3. ¿Qué efecto tiene la IA?	11
3.1 Los efectos de la IA en el trabajo y vida futura y sus implicaciones éticas	11
3.2 Los efectos de la IA en el trabajo futuro.....	11
3.3 Habilidades y actitudes éticas para preparar a los humanos para vivir y trabajar con IA....	12
4. Evaluación	14
4.1 Rúbrica de evaluación para los monitores	15

ARTIN FUTURE – ÉTICA EN LA INTELIGENCIA ARTIFICIAL

A. Introducción

Sobre esta herramienta

La inteligencia artificial (IA) no se trata solo de una palabra de moda, sino de una parte esencial de nuestras vidas. Cuando buscamos algo en internet, la inteligencia artificial nos proporciona resultados relevantes de nuestra búsqueda basándose en nuestras preferencias y actividad previa. La IA nos suele recomendar qué películas ver o qué comprar después. Incluso nos ayuda a comunicarnos entre nosotros al traducir frases y textos de una lengua a otra. Aunque facilita nuestras vidas, muchos temen que nos quite el trabajo, nos discrimine o amenace nuestro derecho a la intimidad. Algunas situaciones en el pasado han demostrado que existe la posibilidad de que se vuelva una amenaza y que debe aumentar la concienciación en torno a la ética de la IA. Con esta herramienta, exploraremos lo bueno y lo malo de la IA y aprenderemos más sobre los principales problemas éticos de la IA.

Resumen

En este documento presentamos una visión general del material de formación 'ArtIn Future – Ética en la inteligencia artificial' para acabar con los mitos de la IA, mostrar las diferencias de la IA y los principios éticos y potenciar el pensamiento crítico como en un formato de herramienta. Contiene información principal sobre los destinatarios del proyecto, las metas y los objetivos, así como información didáctica sobre los métodos de aprendizaje, recursos y actividades.

Este documento debería servir de apoyo a los monitores para aplicar el material didáctico que lo complementa. Está diseñado de modo que ofrezca adaptaciones rápidas a grupos individuales de alumnos, o incluso de diferentes niveles de conocimientos/competencias. Forma parte de los tres resultados principales del producto y puede usarse de manera individual o junto a los otros dos: ArtIn Future Digital skills, un curso de formación experimental para desarrollar las habilidades digitales de los monitores que trabajan, particularmente con alumnos jóvenes (adultos) que exploran la IA *débil* (algoritmos que ya hoy leen patrones de conjuntos de datos que los humanos no podrían reconocer) y/o el ArtIn Future - AI in the Business world, un centro de investigación sobre educación abierta que presenta casos prácticos de IA de empresas, productos diseñados gracias a la IA, historias de éxito de emprendedoras (mujeres) en este campo para ayudar a entender mejor cómo funciona la IA y qué competencias son necesarias para estas tareas.

BEST Institut für berufsbezogene Weiterbildung und Personaltraining GmbH ha creado la siguiente guía. Puede ser utilizada para uso no comercial por cualquier parte interesada haciendo referencia al proyecto (nombre y financiación) y autores. Cualquier otro uso requiere un acuerdo previo por escrito con los autores. Sitio web del proyecto: www.ai-future-project.eu Facebook page: facebook.com/artin-future
Más información sobre el proyecto también aquí:

B. Información de la herramienta

Destinatarios (directos) del proyecto: personal docente y de formación

Beneficiarios finales: estudiantes de todas las edades, escuelas y empresarios que usan la IA en su trabajo

Metas pedagógicas:

- Queremos acabar con los mitos de la IA para que los alumnos y docentes sepan qué es. Introducimos la problemática ética en el campo de la IA y les equipamos con las habilidades y conocimientos necesarios para capacitarles en la evaluación ética de la inteligencia artificial.

Objetivos de enseñanza:

- Trás finalizar las actividades de aprendizaje, los participantes serán capaces de demostrar que comprenden las éticas y la IA y analizar cuestiones relacionadas con la IA desde una perspectiva ética.

Objetivos de aprendizaje:

- Los alumnos adquieren conocimientos sobre los fundamentos de la IA, es decir, terminología necesaria, tipos de IA, etc. Son conscientes de que la IA está en todos lados y de cómo esta influencia nuestra vida tanto de manera positiva como negativa; señalamos la relevancia de la ética en este ámbito. Finalmente, son capaces de evaluar de manera crítica las implicaciones éticas del uso de la IA en la vida diaria. Más precisamente, les ayudamos a calibrar las experiencias en línea basadas en la IA (como buscar en internet, redes sociales, tiendas y servicios) y las experiencias que no están en línea.

Conocimientos	Habilidades	Responsabilidad y autonomía
Conocimientos amplios, especializados, factuales y teóricos en la ética y en la IA y conciencia de los límites de ese conocimiento	Amplio rango de habilidades cognitivas necesarias para desarrollar soluciones para el uso ético de la IA	Gestiona y supervisa en contextos de actividades de estudio y trabajo donde hay cambios imprevisibles; revisar y desarrollar el rendimiento de uno mismo y de los otros

Conjunto de habilidades necesarias para los monitores: un conjunto de habilidades técnicas e interpersonales para transferir el conocimiento de este tema a los beneficiarios del proyecto de distintos grupos de edad. Además de tener conocimientos técnicos y esenciales sobre la IA. Los monitores deben tener una buena capacidad de comunicación y de organización que les permite encontrar formas óptimas de explorar la ética de la IA con los alumnos. Se espera que los monitores tengan buenas habilidades de pensamiento crítico para promover el pensamiento crítico entre los alumnos.

ARTIN FUTURE – ÉTICA EN LA INTELIGENCIA ARTIFICIAL

Metodología de enseñanza sugerida:

- Aprendizaje combinado; aula invertida

- **Aprendizaje combinado en pocas palabras**

Es un tipo de enseñanza y aprendizaje que convierte a la clase en una combinación de sesiones presenciales y en línea.

En las partes presenciales, los monitores interactúan activamente con los alumnos y les animan a participar en los debates y a hacer las actividades prácticas. En las partes en línea, usan diferentes materiales multimedia para apoyar a los alumnos en el proceso de aprendizaje para hacer que estas fases sean tan interactivas como las sesiones presenciales.

- **Aula invertida en pocas palabras**

Es un tipo de aprendizaje combinado que se centra más en los propios alumnos. Ellos son quienes construyen activamente su cuerpo de conocimiento sobre un tema concreto y evalúan su progreso en clase. Los docentes solo facilitan la construcción de este conocimiento y trabaja con los alumnos en las preguntas y problemas que puedan aparecer. Aquí, se anima a los alumnos a tener una sensación de posesión sobre los conocimientos que consiguen a medida que leen/ reciben material primero en casa como vídeos didácticos, investigaciones en línea o similares. De esta manera van a clase listos para debatir lo que han aprendido antes y para explorar casos/ prácticas específicas basadas en el conocimiento más teórico que han adquirido antes o para trabajar en laboratorios con sus compañeros para resolver problemas específicos sobre el asunto concreto sobre el que están reflexionando.

El aula invertida garantiza que todos los participantes aprenden en un ambiente seguro, confiable y libre y que se les demuestra por qué todo lo que aprenden importa. Una vez que son conscientes de la importancia de uno de los temas, se interesan más y se involucran más en el proceso de aprendizaje.

- Métodos de enseñanza:

Discutir sobre ética requiere reflexionar, debatir y analizar las cuestiones importantes y plantear preguntas relevantes. Para mejorar estas habilidades cognitivas, sugerimos al monitor que puede organizar las siguientes actividades de enseñanza:

 - Conversaciones abiertas en clase, es decir, el monitor comienza una conversación y observa como los alumnos interactúan entre ellos a partir de ese impulso.
 - El debate da a los alumnos la oportunidad de presentar los aspectos negativos y positivos de una cuestión
 - El organizador avanzado es bueno para visualizar el tema y la cuestión principal. Puede encontrar más información al respecto aquí:
http://www.projectlearnnet.org/tutorials/advance_organizers.html

ARTIN FUTURE – ÉTICA EN LA INTELIGENCIA ARTIFICIAL

- La técnica de la pecera, una buena técnica para que los alumnos escuchen y aprendan de lo que sus compañeros tienen que decir.
Para aprender más de este método, entre en el enlace siguiente:
<https://www.facinghistory.org/resource-library/teaching-strategies/fishbowl>
- La mantelería es otro ejercicio que permite que los alumnos visualicen y expresen sus pensamientos y opiniones por escrito.
En la página siguiente un vídeo y un documento en pdf les proporcionará más información sobre este método:
https://jct.ie/jcis/strategies_result.php?strategy_id=18
- Técnica del rompecabezas. Permite que toda la clase llegue a la solución de un problema; cada equipo trabaja en una parte de la misma. Cuando la presentan juntos llegan a la conclusión de la misma. En el siguiente enlace hay más información sobre esta técnica: <https://www.teachhub.com/teaching-strategies/2016/10/the-jigsaw-method-teaching-strategy/>

Los métodos pedagógicos mencionados previamente sirven tanto para enseñar a los jóvenes como a los adultos.

Duración:

- Cada lección durará unas 3 horas (cada tema cuenta con 3 clases, se estima que cada una dure unos 45-50 minutos, además se añadirá tiempo para las tareas y el aprendizaje); hay 3 lecciones previstos

Recursos educativos:

- PC/tableta/portátil/móvil, proyector, pizarra/pizarra en línea, rotuladores, presentaciones en PowerPoint de ArtIn Future

Materiales didácticos:

- Presentaciones en PowerPoint de ArtIn Future, contenido audiovisual relevante (los enlaces para cada vídeo aparecerán al final de cada tema)
- Los vídeos que aparecen más abajo hablan brevemente sobre los principales asuntos del tema o mostraran brevemente ejemplos relevantes. Los monitores pueden usar los vídeos tanto como un recordatorio como al principio del tema, por ejemplo, para introducir el tema y/o refrescar la memoria de los alumnos, lo que debería basarse también en su grupo de edad y en la dinámica del grupo.
- Los alumnos y monitores utilizarán las presentaciones de ArtIn Future. Los alumnos recibirán versiones sin los ejercicios que se harán en clase. Los monitores podrán ver las diapositivas ocultas y usar toda la presentación en clase.
- Folletos para los monitores (*¿Qué es la IA?*)
- Folletos para alumnos con preguntas para acabar cada tema y un folleto de ejercicios de la unidad 1. 2. Nota: todo el material didáctico desarrollado puede entenderse como una muestra inicial que más adelante puede usarse o adaptarse para trabajar con alumnos de otros grupos de edad.

Assessment:

- Los educadores evalúan a los alumnos según su participación y trabajo individual. Para la evaluación final, los alumnos deben elegir un dilema ético relacionado con alguno de los temas tratados en el curso. Presentarán su análisis del problema al instructor. (Duración: 1 hora).

ARTIN FUTURE – ÉTICA EN LA INTELIGENCIA ARTIFICIAL

C. Información sobre cada unidad

En las páginas siguientes, presentamos información detallada sobre cada unidad. De este modo, los monitores del aprendizaje pueden cubrir la lección y organizar una clase. Comentamos los objetivos de enseñanza y aprendizaje, las actividades, materiales de aprendizaje y duración (mínimo sugerido). Al final de cada tema, hay una sección de *Lecturas adicionales* donde los monitores del aprendizaje pueden obtener más información sobre el tema.

1. ¿Qué es la IA? (AI débil, AI fuerte, superinteligencia artificial SIA)

1.1 Introducción a la inteligencia artificial

Los participantes del curso aprenderán lo básico sobre la inteligencia artificial, es decir, cómo se define, qué tipos de IA existen y la terminología en torno al concepto de IA.

Objetivos pedagógicos:

- Discutir con los estudiantes lo fundamental de la IA
- Explorar las diferencias entre la IA débil, IA fuerte y la superinteligencia artificial

Objetivos de aprendizaje:

- Tener una base teórica sólida de la IA
- Tener la habilidad de reconocer distintas formas de IA

Metodología de enseñanza:

- Aprendizaje combinado
- Aula invertida

Actividades de enseñanza:

- Debate/conversación abierta y colaboración

Materiales didácticos

- Presentación en PowerPoint de ArtIn Future, contenido audiovisual relevante
- Folletos para los educadores

Duración:

- 45 minutos

Lecturas adicionales:

Explicación detallada de la IA en la Enciclopedia de Filosofía de Stanford:

<https://plato.stanford.edu/entries/artificial-intelligence/>

La Brookings Institution sobre la IA: <https://www.brookings.edu/research/what-is-artificial-intelligence/>

Un artículo sobre las principales diferencias entre IA débil, fuerte y superinteligencia:

<https://medium.com/mapping-out-2050/distinguishing-between-narrow-ai-general-ai-and-super-ai-a4bc44172e22>

Definiciones de IA débil y fuerte de Technopedia:

<https://www.techopedia.com/definition/32874/narrow-artificial-intelligence-narrow-ai>

<https://www.techopedia.com/definition/31622/strong-artificial-intelligence-strong-ai>

Ejemplos típicos del uso de la IA en el día a día: https://medium.com/@the_manifest/16-examples-of-artificial-intelligence-ai-in-your-everyday-life-655b2e6a49de

Archivo de la revista Wired con historias relacionadas con la IA:

<https://www.wired.co.uk/search?q=AI&sort=score+desc>

Un vídeo de YouTube sobre la IA: <https://www.youtube.com/watch?v=nASDYRkbQIY>

Una charla Ted sobre cómo empoderarse, no sobrepasarse, con la IA:

https://www.ted.com/talks/max_tegmark_how_to_get_empowered_not_overpowered_by_ai/transcript#t-220909

ARTIN FUTURE – ÉTICA EN LA INTELIGENCIA ARTIFICIAL

1.2 IA más detalladamente

En esta parte del tema, los participantes se tomarán su tiempo para investigar y escribir las definiciones de algunos de los términos más importantes de la IA.

Objetivos pedagógicos:

- Ampliar los conocimientos sobre la IA

Objetivos de aprendizaje:

- Crear una base teórica firme sobre la IA

Metodología de enseñanza:

- Aprendizaje combinado
- Aula invertida

Actividades de enseñanza:

- Investigación, organizador de avances, debate

Materiales didácticos:

- Presentación en PowerPoint de ArtIn Future, folleto de ArtIn Future

Duración:

- 1,5 horas

Lecturas adicionales:

Glosario de IA del Consejo de Europa: <https://www.coe.int/en/web/artificial-intelligence/glossary>

Qué es el aprendizaje automático de IBM: <https://www.ibm.com/cloud/learn/machine-learning>

Artículo del MIT sobre redes neuronales y aprendizaje profundo:

<https://news.mit.edu/2017/explained-neural-networks-deep-learning-0414>

Aprendizaje automático supervisado y sin supervisión:

<https://www.bmc.com/blogs/supervised-vs-unsupervised-machine-learning/>

Aprendizaje por refuerzo: <https://towardsdatascience.com/the-ultimate-beginners-guide-to-reinforcement-learning-588c071af1ec>

1.3 Principios éticos y su relevancia para las TIC e IA

Los participantes debaten sobre el progreso y la presencia de la IA en la vida cotidiana y las implicaciones éticas de su uso.

Objetivos pedagógicos:

- Introducir los problemas éticos en el ámbito de la IA

Objetivos de aprendizaje:

- Entender la necesidad de la ética de la IA
- Incorporar el marco ético en los debates sobre la IA
- Analizar la aplicación de la inteligencia artificial desde una perspectiva ética

Metología de enseñanza:

- Aprendizaje combinado
- Aula invertida

Métodos de enseñanza:

- Debate y colaboración, técnica del rompecabezas

ARTIN FUTURE – ÉTICA EN LA INTELIGENCIA ARTIFICIAL

Materiales didácticos:

- Presentación en PowerPoint de ArtIn Future, contenido audiovisual relevante

Duración:

- 45 minutos

Lecturas adicionales:

La revista Wired hace un breve resumen sobre la problemática de la ética de la IA: <https://www.wired.co.uk/article/ai-ethics-law>

Planteamiento de la UE sobre la IA: <https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence>

Directrices éticas para una IA fiable propuesta por la UE:

<https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>

Principales principios éticos de la IA de la OCDE:

<https://www.oecd.ai/ai-principles>

Directrices del Instituto Alan Turing para una IA responsable y

ética: [https://www.turing.ac.uk/sites/default/files/2019-](https://www.turing.ac.uk/sites/default/files/2019-06/understanding_artificial_intelligence_ethics_and_safety.pdf)

[06/understanding_artificial_intelligence_ethics_and_safety.pdf](https://www.turing.ac.uk/sites/default/files/2019-06/understanding_artificial_intelligence_ethics_and_safety.pdf)

Vídeo de YouTube sobre ética: <https://www.youtube.com/watch?v=u399XmkjeXo>

Charla TEDx sobre la IA y la ética: <https://www.youtube.com/watch?v=HSsQApXQGsl>

Vídeo de Accenture sobre la ética y la IA: <https://www.youtube.com/watch?v=x9gan8vOBJ8>

ARTIN FUTURE – ÉTICA EN LA INTELIGENCIA ARTIFICIAL

2. ¿Qué hace la IA?

2.1 El potencial de la IA

Aquí, los participantes aprenden sobre el impacto positivo de la IA y los beneficios que puede ofrecer.

Objetivos pedagógicos:

- Presentar el potencial de la IA desde una perspectiva económica, técnica y social

Objetivos de aprendizaje:

- Reconocer los beneficios de la IA en el día a día

Metodología de enseñanza:

- Aprendizaje combinado
- Aula invertida

Actividades de enseñanza:

- Debate, colaboración

Materiales didácticos:

- Presentaciones en PowerPoint de ArtIn Future, contenido audiovisual relevante

Duración:

- 45 minutos

Lecturas adicionales:

Artículo de la UE sobre los beneficios de la IA: <https://digital-strategy.ec.europa.eu/en/library/artificial-intelligence-real-benefits>

Breve resumen sobre los beneficios potenciales de la IA para la economía y la sociedad: <https://hackr.io/blog/benefits-of-artificial-intelligence>

Vídeo de YouTube sobre los 10 beneficios de la IA: <https://www.youtube.com/watch?v=masnR4-vt3M>

2.2 Los riesgos de la IA

En esta parte del tema, se habla sobre el impacto negativo y los riesgos de usar la IA.

Objetivos pedagógicos:

- Presentar los riesgos de la IA

Objetivos de aprendizaje:

- Reconocer las desventajas de la IA en la vida diaria

Metodología de enseñanza:

- Aprendizaje combinado
- Aula invertida

Actividades de enseñanza

- Facilitación del aprendizaje, colaboración

Materiales didácticos:

- Presentación en PowerPoint de ArtIn Future, contenido audiovisual relevante

Duración:

- 45 minutos

ARTIN FUTURE – ÉTICA EN LA INTELIGENCIA ARTIFICIAL

Lecturas adicionales:

Resumen del Foro Económico Mundial sobre los riesgos de la IA::

<https://www.weforum.org/agenda/2016/10/top-10-ethical-issues-in-artificial-intelligence/>

Artículo de Future of Life sobre los riesgos de los mitos sobre la IA:

<https://futureoflife.org/background/benefits-risks-of-artificial-intelligence/>

Base de datos de incidentes de IA que recoge todos los percances que hay entorno a la IA:

<https://incidentdatabase.ai/summaries/incidents>

Vídeo de YouTube sobre los 10 riesgos de la IA: <https://www.youtube.com/watch?v=1oeoosMrJz4>

2.3 Estudio de caso del potencial y los riesgos de la IA

En esta parte del tema basada en la práctica, los alumnos tendrán la oportunidad de elegir una tecnología de IA y analizarla. Se espera que cubran tanto el potencial como los riesgos de la IA elegida y usar el marco de trabajo de los temas 2.1. y 2.2. Más concretamente, deben presentar el potencial y los riesgos económicos, técnicos y sociales de la IA.

Objetivos pedagógicos:

- Aplicar los conocimientos teóricos en la práctica

Objetivos de aprendizaje:

- Reconocer las ventajas y desventajas de las diferentes tecnologías de la IA

Metodología de enseñanza:

- Aprendizaje combinado
- Aula invertida

Actividades de enseñanza:

- Facilitación del aprendizaje, trabajo en grupo, debate, reflexión

Actividades didácticas:

- Estudio de caso

Duración:

- 1,5 horas

ARTIN FUTURE – ÉTICA EN LA INTELIGENCIA ARTIFICIAL

3. ¿Qué impacto tiene la IA?

3.1 Los efectos de la IA en el trabajo y vida futura y sus implicaciones éticas

Después de debatir, en las unidades previas, los aspectos positivos y negativos de la IA, los alumnos pueden centrarse, en los temas siguientes, en las implicaciones éticas de la IA en el trabajo y en la vida.

Objetivos pedagógicos:

- Resumir las tendencias y previsiones en el campo de la IA
- Presentar las oportunidades e inconvenientes de la IA en la vida

Objetivos de enseñanza:

- Desarrollar las habilidades de pensamiento crítico basadas en soluciones de IA
- Adquirir conocimiento sobre la IA en su (futura) vida profesional

Metodología de enseñanza:

- Aprendizaje combinado
- Aula invertida

Actividades de enseñanza:

- Discusión, debate, reflexión

Materiales didácticos:

- Presentación en PowerPoint de ArtIn Future, contenido audiovisual relevante

Duración:

- 45 minutos

Lecturas adicionales:

Un artículo sobre el impacto de la IA en el día a día:

<https://towardsdatascience.com/how-artificial-intelligence-is-impacting-our-everyday-lives-eae3b63379e1>

Una lista de ejemplos de cómo usar la IA en el día a día:

<https://www.forbes.com/sites/bernardmarr/2019/12/16/the-10-best-examples-of-how-ai-is-already-used-in-our-everyday-life/>

Un artículo sobre el papel de la IA en el lugar de trabajo: <https://fowmedia.com/future-work-ai-entered-workplace/>

An article from Forbes about AI and its effects on employees' experience:

<https://www.forbes.com/sites/bernardmarr/2019/05/29/artificial-intelligence-in-the-workplace-how-ai-is-transforming-your-employee-experience/>

Un vídeo de YouTube sobre cómo usamos la IA en el día a día:

<https://www.youtube.com/watch?v=MpR6JZdQ4B0>

Un vídeo de YouTube sobre el impacto de la IA en nuestras vidas:

<https://www.youtube.com/watch?v=-aNW08ahM68>

3.2 Los efectos de la IA en el trabajo futuro

Después de conocer los efectos de la IA en el día a día, los alumnos tendrán que pensar en el futuro y en qué efectos tendrá la IA cuando empiecen a trabajar. En otro ejercicio prácticos, los alumnos debatirán cómo afectará la IA a la mano de obra del futuro. .

Objetivos pedagógicos:

- Aplicar el conocimiento teórico de la ética en la IA en práctica

ARTIN FUTURE – ÉTICA EN LA INTELIGENCIA ARTIFICIAL

Objetivos de aprendizaje:

- Tener la habilidad de predecir el impacto de la IA en el trabajo y vida laboral

Metodología de enseñanza:

- Aprendizaje compuesto
- Aula invertida

Actividades de enseñanza:

- Debate en pecera, reflexión

Materiales didácticos:

- Presentación en PowerPoint de ArtIn Future

Duración:

- 1,5 horas

3.3 Habilidades y actitudes éticas para preparar a las personas para vivir y trabajar con tecnología de IA

En esta parte de la unidad, el grupo desarrollará su propia guía ética para la IA.

Objetivos pedagógicos:

- Resumir los principios éticos relevantes para la IA que pueden beneficiar a la sociedad

Objetivos de aprendizaje:

- Comprender los principios relevantes relacionados con la ética y la IA
- Desarrollar las propias actitudes éticas hacia la IA

Metodología de enseñanza:

- Aprendizaje combinado
- Aula invertida

Actividades de enseñanza:

- Facilitación del aprendizaje, mantelería, reflexión

Materiales didácticos:

- Presentación en PowerPoint de ArtIn Future

Duración:

- 45 minutos

Lecturas adicionales:

Un artículo de Harvard Data Science Review sobre el marco de trabajo ético para la IA:

<https://hdr.mitpress.mit.edu/pub/l0jsh9d1/release/7>

Marco de trabajo ético para la IA de Asilomar: <https://futureoflife.org/ai-principles/>

Reflejos de IEEE para una IA ética: https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ead1e_general_principles.pdf?utm_medium=undefined&utm_source=undefined&utm_campaign=undefined&utm_content=undefined&utm_term=undefined

Declaración de la UE sobre la IA y una propuesta de marco ético:

<https://op.europa.eu/en/publication-detail/-/publication/dfebe62e-4ce9-11e8-be1d-01aa75ed71a1/language-en/format-PDF/source-78120382>

Guías de la UE para la IA ética: [https://digital-](https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai)

[strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai](https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai)

ARTIN FUTURE – ETHICS IN ARTIFICIAL INTELLIGENCE

4. Evaluación

Los docentes deben evaluar a los alumnos según su participación y trabajo individual como actividad continua. Las preguntas planteadas en las presentaciones de ArtInFuture pueden ser de ayuda para ver el progreso del aprendizaje relacionado con la comprensión de la IA y los principios éticos que deberían considerar a la hora de aplicar la IA. También se proporciona un modelo de folleto que cubre los tres temas que pueden usarse, puede preguntarse al final de cada tema o entregarse a los alumnos poco a poco.

Para su evaluación final, los alumnos tendrán que dar su opinión sobre un dilema ético. Un dilema ético es una situación que tiene muchas 'soluciones', las opciones que barajea la persona son, normalmente, más difíciles y no solucionan el problema totalmente. Piense en el famoso dilema del tranvía: en dos vías hay personas atadas. En la primera vía, hay una persona. En la otra, cinco. Ninguna se puede mover y la persona tiene que elegir a quién sacrifica con el tranvía. ¿Debe el tranvía matar a cinco personas para matar a uno o viceversa?

En nuestro caso, se espera que los alumnos resuelvan dilemas éticos similares relacionados con la IA. Al igual que muestra el ejemplo anterior, el tranvía puede sustituirse por un coche autoconducido. Presente a los alumnos un caso en el que el coche falla de repente y el alumno tiene que elegir a quien atropellar. Por ejemplo, pueden elegir entre un anciano que está cruzando en rojo y un niño que está cruzando en verde. El alumno debe justificar su respuesta. ¿Sus sugerencias cambiarían si se tratase de una mujer joven con un bebé y un grupo de gente sin definir? Otro ejemplo que puede no parecer tan dramático es también ver la ética en situaciones, por ejemplo, donde se tiene una conversación con bots que imitan a sus seres queridos fallecidos (en línea/ con mensajes a través del móvil...) o en páginas de citas en las que bots llevan la conversación sin decir a sus usuarios que no son humanos.

La página web de MIT Moral Machine ofrece numerosos ejemplos que los monitores pueden presentar a los alumnos para evaluarlos. Aquí, los docentes también pueden diseñar sus propios dilemas de IA. Se puede encontrar más información aquí: <https://www.moralmachine.net/>

La UNESCO también ofrece dilemas relevantes éticos que pueden ayudar a los docentes en la evaluación. Se puede encontrar más información aquí: <https://en.unesco.org/artificial-intelligence/ethics/cases>

La evaluación final se puede realizar como un trabajo en grupo o de manera individual. Sería importante juzgar el resultado en función de la capacidad de implementar y aplicar lo que se ha aprendido en relación con la ética y la IA en los tres temas. Se sugiere una rúbrica de evaluación para ayudar con la tarea evaluativa.

Duración de la evaluación: 1 hora

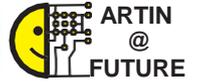
ARTIN FUTURE – ÉTICA EN LA INTELIGENCIA ARTIFICIAL

4.1 Rúbrica de evaluación para los monitores

A continuación encontrará una rúbrica que le ayudará a llevar a cabo la evaluación de la tarea. La rúbrica puede modificarse y usarse no solo para evaluar esta tarea, sino para evaluar el desempeño general del alumno.

Pasa	Satisfactorio	Bueno	Muy bueno	Excelente
El/la alumno/a demuestra conocimientos básicos de la IA y de ética. Responde al dilema sin explicarse detalladamente.	La comprensión del dilema es alta. Entiende los riesgos que puede conllevar la IA y reconoce la necesidad de la ética de la IA, pero dentro de unos límites.	El/la alumno/a es capaz de reconocer ambas soluciones para el dilema al que se enfrentan y respuestas de manera más clara y reconoce los riesgos de elegir una opción.	El/la alumno/a no solo demuestra que las opciones son opuestas, también reconoce que la ética discrepa a veces y que se necesitan principios más claros.	El/la alumno/a presenta en su respuesta la naturaleza conflictiva de los principios éticos y de la IA, indica que estas deberían crearse para beneficiar a todos los humanos y no solo a ciertos grupos y sugiere nuevas reglas o presenta nuevas formas de repensar las reglas.

Co-funded by the
Erasmus+ Programme
of the European Union



ARTIN FUTURE



TECHNISCHE
UNIVERSITÄT
DRESDEN



DAUGAVPILS TEHNIKUMS



Este proyecto ha sido financiado con el apoyo de la Comisión Europea en el marco del programa Erasmus+. Esta publicación refleja únicamente la visión del autor. La Comisión no se hace responsable del uso que pueda hacerse de la información contenida en ella. (2020-1-DE02-KA226-VET-00813)