

Part of

Educación Digital Crítica en
Adolescentes: Impacto de los sesgos
en el uso de herramientas digitales
(EDICA)

Dataset

Educación Digital Crítica en
Adolescentes: Impacto de los
sesgos en el uso de
herramientas digitales (EDICA)

Description

Educación Digital Crítica en Adolescentes: Impacto de los sesgos en el uso de herramientas digitales (EDICA)

Researchers

Juana M. Ortega-Tudela, Mercedes Llorent Vaquero

Funder

--

Grant

--

Dataset Description

1.1 Resumen de los datos

1.1.1 Indique información general sobre el proyecto: Título, descripción breve del mismo y duración.

Título del Proyecto:

Educación Digital Crítica en Adolescentes: Impacto de los sesgos en el uso de herramientas digitales (EDICA).

Descripción Breve:

El proyecto se centra en analizar cómo los adolescentes perciben los sesgos presentes en la inteligencia artificial (IA) y las redes sociales (RRSS), y en desarrollar un programa educativo que los capacite para identificar y mitigar dichos sesgos. El objetivo es fomentar una ciudadanía digital crítica, inclusiva y

responsable, abordando problemáticas como la polarización, la perpetuación de estereotipos de género y la desinformación. Se prioriza la educación digital crítica para adolescentes, quienes son especialmente vulnerables a los efectos negativos de estos sesgos.

Duración:

El proyecto se estructura en seis fases a lo largo de tres años, incluyendo un diagnóstico inicial, diseño y validación de la formación, implementación en centros educativos, evaluación de impacto y difusión de resultados.

1.1.2 Indique quiénes son los financiadores del proyecto.

Se presenta a la Convocatoria de Proyectos de Generación del conocimiento 2024

1.1.3 Indique los objetivos de los datos generados y / o recolectados y su relación con los objetivos del proyecto.

Objetivos de los Datos Recolectados:

Identificar y caracterizar los sesgos comunes en IA y RRSS:

Los datos recolectados permitirán construir un mapa de los sesgos más recurrentes (como sesgos de confirmación, de género o de exclusión) y generar una base de datos con ejemplos concretos.

Analizar las interacciones de los adolescentes con IA y RRSS:

Mediante cuestionarios y análisis de hábitos de consumo, se busca comprender el perfil de usuario adolescente, las plataformas que utilizan, los contenidos más frecuentes y los creadores que siguen.

Evaluar la percepción de los adolescentes sobre los sesgos:

Investigar cómo perciben los adolescentes la influencia de los sesgos en sus decisiones y opiniones, incluyendo el impacto del género y la diversidad en los contenidos que consumen.

Determinar el nivel de conocimiento de los adolescentes sobre los sesgos:

Se busca evaluar su comprensión acerca de los sesgos algorítmicos, su impacto y los mecanismos para mitigarlos.

Evaluar la capacidad para identificar y mitigar los sesgos:

Analizar las habilidades críticas de los adolescentes para reconocer y contrarrestar los efectos negativos de los sesgos en su entorno digital.

1.1.4 Indique el tipo y el formato de los datos que se generarán / recolectarán.

Tipo de Datos:

Cuantitativos: Recopilados a través del cuestionario CESIA, que mide hábitos, percepciones y conocimientos de los adolescentes sobre sesgos en IA y redes sociales.

Cualitativos: Datos obtenidos de transcripciones de sesiones de reflexión-acción, discusiones en grupos focales y entrevistas.

Visuales: Material multimedia como recursos educativos (videos, dinámicas, etc.).

Formatos:

Datos Cuantitativos:

Hojas de cálculo en formatos abiertos como CSV o XLSX.

Datos Cualitativos:

Transcripciones textuales en formatos TXT o PDF.

Visuales y Multimedia:

Imágenes en formatos PNG o JPEG.

Videos en formatos MP4.

Recursos educativos en formatos abiertos compatibles con la web.

Consideraciones:

Los datos serán tratados según los principios FAIR (localizables, accesibles, interoperables y reutilizables).

Formatos abiertos serán priorizados para garantizar la preservación y accesibilidad en repositorios digitales, como la Nube Europea de Ciencia Abierta (EOSC) y el Repositorio RUJA

1.1.5 Indique el origen de los datos que se generarán, recolectarán o se reutilizarán de otras fuentes: encuestas, sensores térmicos, trabajos previos, bases de datos, datos de otros proyectos,...

1. Datos Generados:

Encuestas: Se desarrollará y aplicará el Cuestionario CESIA (Cuestionario de Evaluación del Impacto de los Sesgos de la IA y las Redes Sociales en la Educación Digital Crítica Adolescente) para recopilar datos sobre hábitos, percepciones y capacidades de los adolescentes respecto a los sesgos digitales.

Sesiones de Reflexión-Acción: Datos cualitativos generados durante las actividades educativas en talleres y auditorías en el marco del Modelo EDICA.

2. Datos Recolectados:

Análisis de Contenido: Revisión de publicaciones, tendencias en redes sociales y ejemplos concretos de sesgos en herramientas digitales (IA y RRSS).

Interacciones Digitales: Datos de los adolescentes sobre plataformas, contenidos y creadores de contenido que siguen, así como patrones de uso.

3. Datos Reutilizados:

Trabajos Previos y Bases de Datos: Uso de literatura científica y revisiones sistemáticas de estudios previos sobre sesgos en IA y RRSS, obtenidos de bases de datos como WoS, Scopus, ERIC y Dialnet.

1.1.6 ¿Cuál es el tamaño esperado de los datos?

--

2.1 Los datos y metadatos se deben localizar fácilmente.

2.1.1 Indique el tipo de identificador persistente que se usará para facilitar la cita del dataset depositado en un repositorio.

El proyecto empleará identificadores persistentes como el DOI (Digital Object Identifier) para los datasets que se depositen en repositorios. Este tipo de identificador es ampliamente utilizado en la comunidad científica y garantiza:

Accesibilidad: Facilita el acceso directo al dataset desde cualquier plataforma.

Citación: Permite a otros investigadores referenciar el dataset correctamente en sus trabajos.

Preservación: Asegura la localización y preservación del dataset a largo plazo.

Los datos estarán disponibles en repositorios reconocidos, como el Repositorio RUJA de la Universidad de Jaén, la Nube Europea de Ciencia Abierta (EOSC) y/o plataformas como OSF (Open Science Framework).

2.1.2 Describa el sistema que se usará para nombrar los archivos y directorios, así como la estructura de los directorios en el caso de haberlos.

Sistema de Nomenclatura:

Se utilizará un sistema descriptivo y estandarizado para nombrar los archivos, que facilite su identificación y organización. La estructura incluirá los siguientes elementos:

Prefijo del proyecto: Identificación del proyecto, por ejemplo, EDICA.

Fase del proyecto: Indicará la fase específica de la que provienen los datos, como Fase1, Fase2, etc.

Tipo de dato: Especificará si los datos son cualitativos (Cual), cuantitativos (Cuant), o multimedia (Mult).

Fecha: Formato YYYYMMDD para reflejar la fecha de creación o recopilación.

Versión: Indicado como v1, v2, etc., para rastrear cambios y actualizaciones.

2.1.3 Una vez publicado el dataset, ¿podrán publicarse nuevas versiones? Si es así, ¿cómo se nombrarán las distintas versiones?

No en principio

2.1.4 Los metadatos ayudan a localizar el dataset a los investigadores que quieran usarlo. ¿Qué esquemas de metadatos se usan en el repositorio en el que se va a depositar el dataset?

--

2.2 Los datos deben de ser accesibles

2.2.1 Indique si los datos finales de la investigación estarán disponibles en acceso abierto y en qué repositorio se depositarán.

Sí, los datos finales de la investigación estarán disponibles en acceso abierto. Se depositarán en el Repositorio Institucional de Producción Científica de la Universidad de Jaén (RUJA), cuyo objetivo es recopilar, gestionar, presentar y difundir en modo de acceso abierto la producción científica e institucional de la UJA, garantizando su preservación.

2.2.2 Si los datos no pudieran estar en acceso abierto, indique la razón, quién y cómo podrá acceder a los datos y en qué repositorio se depositarán.

--

2.2.3 Indique si el repositorio seleccionado es un repositorio de confianza (trustworthy, trusted). ¿Tiene algún certificado que lo acredite (CoreTrustSeal, ISO 1636, ...)?

Repositorio Institucional de la Universidad de Jaén

2.2.4 Indique si tiene un plan de preservación para el mantenimiento de los datos a medio y largo plazo.

Repositorio Universidad de Jaén

2.3 Los datos deben de ser interoperables

2.3.1 Indique qué vocabularios se usarán en la descripción de los datos para materias (Unesco, JEL, MeSH, ...) o identificadores estándar (ORCID o ResearcherID para autores, ROR para instituciones), ...

Se usarán descriptores de datos de UNESCO e identificadores estándar.

2.3.2 Indique, si lo conoce, si el dataset estará disponible desde otros portales científicos generalistas o temáticos (Google Scholar, OpenAIRE, Recolecta, ...).

--

2.4 Los datos deben de ser reutilizables

2.4.1 Indique las condiciones de acceso a los datos y la licencia para el acceso a los mismos en caso de usar una licencia estándar como las Creative Commons.

Se publicará bajo licencia Creative Commons

2.4.2 Indique, si lo conoce, el tiempo mínimo que el dataset estará accesible.

--

2.4.3 ¿Qué software será necesario para visualizar y reutilizar los datos?

Excel

3.1 Asignación de recursos humanos y económicos

3.1.1 Indique el personal (roles) encargado de generar los datos, tratarlos y preparar el dataset final. Indique quién será el encargado de la gestión de datos del proyecto.

Investigadoras principales

3.1.2 Indique el coste de generar y mantener los datos durante la vida del proyecto si los hubiera (costes directos, indirectos, discos externos, contratación personal auxiliar, preparar los datos para que sean FAIR, ...).

--

4.1 Seguridad de los datos

4.1.1 Indique los mecanismos para garantizar la seguridad y acceso al dataset (copias de seguridad, evitar acceso no autorizados, ...) tanto durante su gestión como cuando ya se han depositado los datos en un repositorio.

La gestión de los datos recogidos se realizará siguiendo los principios FAIR, asegurando que sean localizables, accesibles, interoperables y reutilizables. Para ello, se elaborará un Plan de Gestión de Datos (PGD) que detallará las medidas de seguridad y acceso, incluyendo la realización de copias de seguridad periódicas en diferentes ubicaciones y el establecimiento de permisos de acceso diferenciados según los roles de los usuarios. Se considerará el cifrado de los datos para proteger la información sensible y se implementarán protocolos de seguridad informática para prevenir accesos no autorizados. Además, se registrará la actividad de acceso para detectar posibles anomalías. Los datos se depositarán en repositorios confiables como OSF, el TDR institucional de la Universidad de Jaén, la Nube Europea de Ciencia Abierta (EOSC) o el Repositorio RUJA, que ofrezcan medidas de seguridad robustas. Se utilizarán metadatos descriptivos para facilitar la comprensión y el uso adecuado del dataset por parte de los usuarios, siempre cumpliendo con las normativas de protección de datos y garantizando la anonimización de la información para proteger la privacidad de los participantes.

5.1 Aspectos éticos

5.1.1 Explique los problemas éticos o legales que pueden afectar a la recogida e intercambio de datos.

--

5.1.2 ¿Es necesario anonimizar datos (indique cuáles en su caso) o crear un comité de acceso a los datos

Los datos se anonimizarán con códigos

Powered by

