

ISSN: 0210-1696

DOI: <https://doi.org/10.14201/scero.32147>

## ATECA: UNA APLICACIÓN WEB PARA AGILIZAR LAS ADAPTACIONES A LECTURA FÁCIL

### *ATECA: a Web Application to Facilitate Easy-to-Read Adaptations*

Mari Carmen SUÁREZ DE FIGUEROA BAONZA

*Universidad Politécnica de Madrid. Ontology Engineering Group (OEG). España*  
mcsuarez@fi.upm.es

<https://orcid.org/0000-0003-3807-5019>

Paula Marina BLANCO

*Universidad Politécnica de Madrid. Ontology Engineering Group (OEG). España*  
pm.blanco@upm.es

Acher BLANCO

*Universidad Politécnica de Madrid. España*  
acher.bmartin@alumnos.upm.es

Isam DIAB LOZANO

*Universidad Politécnica de Madrid. Ontology Engineering Group (OEG). España*  
isam.diab@upm.es

<https://orcid.org/0000-0002-3967-0672>

María GARCÍA-AGUDO

*Universidad Politécnica de Madrid. España*  
maria.gestepa@alumnos.upm.es

Alejandro MUÑOZ

*Real Academia Española (RAE). España*  
alejandro@munoznavarro.com

Cynthia TORRES

*Universidad Politécnica de Madrid. España*  
cynthia.tcelorio@alumnos.upm.es

Recepción: 6 de agosto

Aceptación: 21 de noviembre

Ediciones Universidad de Salamanca / CC BY-NC-SA  
*Siglo Cero*, vol. 56 (2), 2025, abril-junio, pp. 115-137

RESUMEN: La lectura fácil es un método de creación de documentos accesibles para personas con discapacidad cognitiva (Plena inclusión, s.f., <https://www.plenainclusion.org/discapacidad-intelectual/recurso/lectura-facil/>). Para la producción de estos textos en lectura fácil, los adaptadores corrigen manualmente los textos y los despojan de cualquier estructura complicada que pueda dificultar el proceso de lectura. En un intento de facilitar esta labor, el presente proyecto de investigación pretende desarrollar una herramienta que agilice el proceso de adaptación. De modo que, en colaboración con el Centro Español de Accesibilidad Cognitiva (CEACOG), se ha desarrollado la aplicación web ATECA (Adaptación de TEXTos en CASTellano) para identificar y adaptar cualquier discrepancia según una selección de las pautas de lectura fácil establecidas por la Norma UNE 153101:2018 EX (AENOR, 2018). ATECA es un sistema de apoyo piloto, basado en inteligencia artificial (IA), que se puede incorporar al proceso tradicional de adaptación manual de documentos a la metodología de lectura fácil. Esta aplicación piloto realiza de manera (semi)automática las dos primeras actividades del proceso cíclico de adaptación de documentos a lectura fácil: el análisis de documentos y la sugerencia de transformación de dichos documentos para que cumplan con las pautas de la metodología.

PALABRAS CLAVE: Inteligencia artificial; lectura fácil; adaptación semiautomática; español.

ABSTRACT: Easy Read arranges texts in an accessible manner so that people with cognitive disabilities may have an easier time reading a clear and simple language. Regarding the production of these easy-to-read texts, professional adapters manually correct the texts and strip them of any complicated structures that may difficult the reading process. In an attempt to facilitate this labour, the present research project aims to create a tool that expedites the adapting process, creating accessible and comprehensive documents for readers of any level. Thus, in collaboration with the Centro Español de Accesibilidad Cognitiva (CEACOG), the online service ATECA (Adaptación de TEXTos en CASTellano) was created to identify and adapt any discrepancy with a selection of the guidelines on Spanish Easy Read established by the Norma UNE 153101:2018 EX (AENOR, 2018). ATECA is a pilot support system, based on Artificial Intelligence (AI), that can be introduced into the traditional process of manually adapting documents following the methodology. This pilot application performs in a (semi-)automatic manner the two first activities in the cyclical process of adapting documents to Easy Read: the analysis of documents and the suggestion for how to transform such documents, so that they comply with the guidelines of the methodology.

KEYWORDS: Artificial intelligence; easy-to-read; semi-automatic adaptation; Spanish.

## 1. Introducción

Un 55.4% de la población española tiene dificultades a la hora de leer y comprender cualquier texto en castellano, englobando a personas con discapacidad, personas no nativas y las personas con un dominio limitado del lenguaje. Por ello, se promueve el uso de la metodología de lectura fácil, que contribuye al acceso de información para aquellas personas con dificultades de

comprensión lectora. La Norma UNE 153101:2018 EX (AENOR, 2018) establece las pautas y recomendaciones a seguir para la creación o adaptación de material a lectura fácil. A la hora de seguir esta metodología, se realizan tres labores principales: 1) una evaluación del documento según el cumplimiento de las pautas; 2) la transformación del documento por adaptadores expertos en lectura fácil, corrigiendo los incumplimientos de la metodología; y 3) la validación del producto final por un grupo profesional compuesto por personas con discapacidad cognitiva.

No obstante, la metodología de lectura fácil no es perfecta; las pautas no indican cómo adaptar diferentes tipos de textos concretos, por ejemplo, un texto literario se diferencia bastante de un documento legal; y la adaptación de material se realiza de forma manual, dificultando el uso de esta metodología en gran escala.

La investigación descrita en este artículo se basa en el desarrollo de una aplicación piloto que: a) evalúe textos escritos en castellano con respecto a la metodología de lectura fácil, tal y como se indica en la Norma UNE 153101:2018 EX (AENOR, 2018); y b) sugiera posibles transformaciones en dichos textos para que cumplan con las pautas de lectura fácil. Esta aplicación piloto está destinada a hacer la labor de adaptación a lectura fácil más dinámica. Esta aplicación ha sido el resultado de la colaboración entre un equipo de investigación de la Universidad Politécnica de Madrid (UPM) y un equipo de trabajo del Centro Español de Accesibilidad Cognitiva (CEACOG), un centro público que investiga e impulsa la accesibilidad cognitiva, que depende del Real Patronato sobre Discapacidad y que está gestionado por Plena inclusión.

El objetivo principal de este proyecto de investigación es desarrollar una aplicación web que facilite la creación de documentos más accesibles y comprensibles para lectores de distintos niveles, promoviendo una comunicación efectiva e igualdad de oportunidades en el acceso a la información. Adicionalmente, este proyecto busca demostrar que un sistema de apoyo, basado en inteligencia artificial (IA), para realizar el proceso global de adaptación de materiales a las pautas indicadas en la metodología de lectura fácil, es de utilidad y facilita el trabajo diario a los adaptadores de documentos.

Basándose en esto, esta investigación presenta el desafío de incorporar al proceso tradicional de adaptación manual de documentos a la metodología de lectura fácil, una aplicación piloto que realiza de manera (semi)-automática las dos primeras actividades del proceso cíclico de adaptación de documentos a lectura fácil: el análisis de documentos y la sugerencia de transformación de dichos documentos para que cumplan con las pautas de la metodología.

Las pautas de lectura fácil descritas en la Norma UNE (AENOR, 2018) abarcan varias partes del texto que pueden dificultar la comprensión, como la redacción, la ortografía, el vocabulario y las expresiones, o la organización y estilo del texto, entre otras. La Norma UNE (AENOR, 2018) especifica 130 pautas que deben cumplirse para la creación o adaptación de textos. Como primer paso en la investigación que describe este artículo, se ha decidido trabajar en las siguientes pautas:

- Vocabulario y expresiones:
  - Se deben evitar los superlativos. Es recomendable añadir el adverbio *muy* al adjetivo o al adverbio.

- Se debería evitar el uso de palabras muy largas o que contengan sílabas complejas.
- Se debería evitar el uso de palabras de contenido indeterminado como: *cosa, algo* o *asunto*.
- Se debe utilizar a lo largo del texto la misma palabra para denominar el mismo objeto o referente.
- Frases y palabras:
  - Se debería evitar el uso de conectores complejos entre oraciones, como *por lo tanto, no obstante, por consiguiente* o *sin embargo*.
  - Se debería utilizar el imperativo solo en contextos claros para evitar que se confunda con la tercera persona del singular del presente de indicativo.
  - Se debe evitar la voz pasiva.
- Ortografía:
  - No se deben escribir palabras ni frases con todas sus letras en mayúsculas, excepto cuando se trate de siglas.
  - Se deberían utilizar los dos puntos (:) cuando se introduce una lista que enumera más de tres elementos.

Para permitir alcanzar los objetivos planteados, esta investigación se ha desarrollado en dos fases. La primera fase se ha centrado en: a) la creación de un primer servicio web para la identificación y sugerencia de cambios de las pautas de lectura fácil sobre conectores complejos, mayúsculas y superlativos descritas en la Norma UNE (AENOR, 2018); y b) el desarrollo de la interfaz de usuario para la aplicación web. Durante la segunda fase, se ha desarrollado en más profundidad el servicio web introduciendo e integrando el resto de las pautas seleccionadas. Todo este proceso se ha realizado a partir de una investigación participativa (Cargo y Mercer, 2008) con expertos en la adaptación de lectura fácil en forma de cuestionarios.

Este artículo está estructurado de la siguiente manera: la sección 2 presenta la metodología de desarrollo utilizada para la creación de ATECA (Adaptación de TExtos en CASTellano); la sección 3 muestra los resultados estadísticos de registro y uso de la aplicación; finalmente, en la sección 4 se incluyen las conclusiones y líneas futuras.

## 2. Metodología

El desarrollo de la aplicación web ATECA se ha realizado en las siguientes etapas: 1) establecimiento del protocolo de gestión del piloto; 2) desarrollo de pautas de lectura fácil; y 3) integración de las pautas desarrolladas en el sistema piloto.

En esta sección resumimos las actividades principales del desarrollo de ATECA. La sección 2.1 describe los requisitos que debe cumplir ATECA; mientras que en la sección 2.2 se presenta el diseño de la interfaz de acceso al sistema piloto. La interfaz de usuario de ATECA se detalla en la sección 2.3; y la sección 2.4 describe el diseño de las pautas seleccionadas para incluir en ATECA. Por último, la sección 2.5 presenta la integración de las pautas en el sistema piloto.

## 2.1. Requisitos

El equipo de trabajo del CEACOG estableció como requisito disponer de los siguientes datos de las personas registradas en la aplicación: tipo de uso de la aplicación, sexo, país de residencia, experiencia en la adaptación de textos a lectura fácil, pertenencia a organización del Tercer Sector, ámbito de trabajo de la organización, vínculo con la organización y trabajo en la Administración Pública.

Además, otro requisito fue la recopilación de datos sobre las operaciones realizadas por los usuarios, con detalle de qué pautas eran seleccionadas en cada operación.

## 2.2. Diseño de acceso al piloto

El sistema de acceso al piloto web ATECA se basa en un proceso de registro de usuario y un proceso de inicio de sesión. La Figura 1 presenta la página inicial de registro de usuarios en ATECA.

La Figura 2 muestra la página de inicio de sesión. Se diseñó una primera interfaz de acceso. Ambas interfaces de inicio fueron validadas por el equipo de trabajo del CEACOG.

FIGURA 1. Página de registro de usuarios en ATECA

**ATECA**   
Adaptación de TExtos en CAstellano

**Registro**

Estas son tus credenciales de inicio de sesión. Guárdalas bien, pues no podrás recuperarlas o consultarlas de nuevo.

Nombre de usuario:  
Código de acceso:

País de residencia \* Comunidad autónoma (si aplica)  
<Selecciona una opción> <Selecciona una opción>

Género \*  
<Selecciona una opción>

**Siguiente →**  He guardado las credenciales en un lugar seguro

Salir  easy2read@delicias.dia.fi.upm.es

FIGURA 2. Página de inicio de sesión en ATECA

ATECA  
Adaptación de TExtos en CAstallano

Usuario  
Introduce tu usuario

Código de acceso  
Introduce tu código de acceso

Iniciar sesión

¿Aún no estás registrado? Crear cuenta

POLITÉCNICA

CENTRO ESPAÑOL DE ACCESIBILIDAD COGNITIVA

easy2read@delicias.dia.fi.upm.es

El proceso de registro de usuarios se lleva a cabo a través de un formulario en el que se debe responder a varias preguntas para recoger la información relevante de los usuarios:

- País de residencia y comunidad autónoma (si aplica).
- Género.
- Uso previsto de la aplicación (profesional, personal u otro).
- Experiencia previa en lectura fácil.
- Pertenencia al Tercer Sector, en cuyo caso:
  - Ámbito de trabajo de la organización.
  - Tipos de discapacidad con los que se trabajan (si aplica).
  - Pertenencia a Plena inclusión (si se trabaja con discapacidad intelectual).
  - Vínculo con la organización (profesional, voluntario, usuario u otro).

- De lo contrario (no pertenece al Tercer Sector):
  - Pertenencia a la Administración Pública.
- Nombre de la organización/Administración Pública.

Adicionalmente, al finalizar este proceso, el usuario revisa y acepta las políticas de protección de datos y da su consentimiento para tratar los datos recogidos con fines de investigación.

### 2.3. *Diseño de la interfaz de usuario del piloto*

Partiendo de la interfaz de usuario de FACILE (Suárez-Figueroa *et al.*, 2024) como base, el equipo de trabajo de la UPM ha elaborado un documento con capturas de las diferentes pantallas de la interfaz de usuario web para que adaptadores de lectura fácil pudieran revisar y analizar cualquier aspecto relevante de la interfaz de usuario.

La Figura 3 muestra una captura de pantalla de la página de inicio que proporciona al usuario la posibilidad de 'Identificar Problemáticas' relacionadas con la metodología de lectura fácil.

Por otro lado, la Figura 4 muestra una captura de pantalla de la interfaz de usuario web para realizar la actividad de 'Obtener Sugerencias de Adaptación de un texto' siguiendo la metodología de lectura fácil.

Además, el equipo de trabajo de la UPM ha creado un documento compartido para recopilar las opiniones de los adaptadores sobre la propuesta de la interfaz de usuario.

Una vez los adaptadores de lectura fácil han proporcionado los comentarios pertinentes sobre la interfaz de usuario propuesta, el equipo de trabajo de la UPM ha analizado en detalle todos los comentarios y sugerencias recopilados. A partir de dicho análisis y los tiempos manejados durante el proyecto, se han realizado las siguientes modificaciones en la interfaz de usuario:

- Nombre diferente para el botón de Identificar "problemáticas" (ver Figura 3).
- Uso de colores diferentes para cada funcionalidad (identificar y sugerir adaptación).
- Mejora del contraste entre texto y fondo.
- Incremento del tamaño del texto (en general).
- Eliminación del contador de pautas seleccionadas.

De manera adicional, y como fruto de diferentes conversaciones internas en el equipo de trabajo de la UPM, se ha decidido incluir en la interfaz de usuario cajas de texto con herramientas de formato (negrita, cursiva, tamaño del texto, entre otros) para que el usuario pueda introducir texto con formato, en lugar de las cajas de texto simples que aparecían en la interfaz de usuario propuesta (ver Figura 3 y Figura 4).

La Figura 5 muestra el diseño final de la página de inicio de la interfaz de usuario de ATECA.

La Figura 6 presenta el diseño final de la interfaz de usuario web para realizar la actividad de 'Obtener sugerencias de adaptación de un texto'.

FIGURA 3. Propuesta de página inicial (Identificar problemáticas)

FIGURA 4. Propuesta de página para obtener sugerencias de adaptación

Con el objetivo de explicar a los adaptadores usuarios de ATECA cómo funciona la aplicación a nivel genérico, se decide que ATECA incluya una sección de documentación. La Figura 7 muestra el diseño de la página de ayuda que contiene: 1) ejemplos de cómo usar la aplicación; 2) mejoras y actualizaciones de ATECA; y 3) explicaciones sobre cómo funciona cada una de las pautas incorporadas en ATECA.

#### 2.4. Diseño de las pautas seleccionadas

Para poder automatizar las pautas seleccionadas (ver listado en sección 1.2), se ha realizado un análisis de viabilidad con el objetivo de investigar si es factible

FIGURA 5. Nueva página inicial (Identificar incumplimientos)

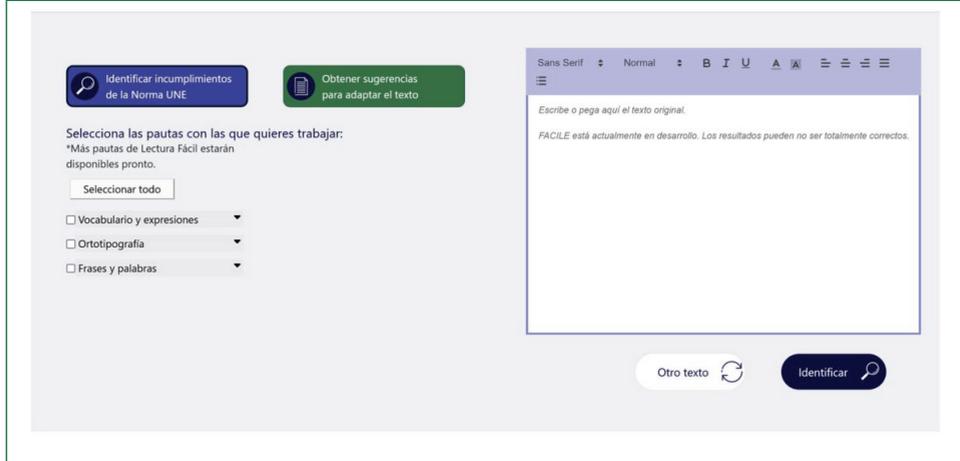
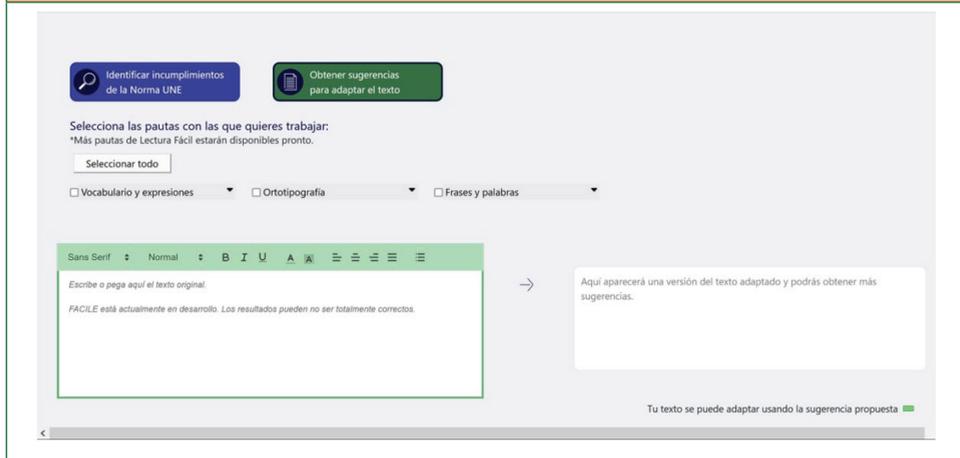


FIGURA 6. Nueva página para obtener sugerencias de adaptación



dicha automatización. Para llevar a cabo este análisis se ha seguido un enfoque incremental en el que se han diseñado y desarrollado diferentes métodos<sup>1</sup> basados en

<sup>1</sup> Un método (IEEE, 1990b) es un conjunto de “procesos o procedimientos ordenados utilizados en la ingeniería de un producto o en la prestación de un servicio”.

FIGURA 7. Página de ayuda de ATECA



técnicas<sup>2</sup> de IA para la automatización de: a) la identificación de las problemáticas mencionadas en las pautas seleccionadas; y b) la propuesta de sugerencias de adaptación de textos en castellano para que las pautas identificadas se cumplan. Este proceso incremental ha comenzado con el diseño y el desarrollo de los métodos utilizando técnicas de la llamada IA simbólica (Russell y Norvig, 2021). Concretamente, se han realizado, en primer lugar, pruebas basadas en sistemas de producción; si el uso de dicha técnica no era efectivo se realizaban pruebas con otras técnicas de manera individual. Si el uso individual de técnicas de IA simbólica no daba buenos resultados, se realizaban pruebas combinando diferentes técnicas.

En las siguientes secciones, se describen los métodos diseñados y desarrollados para automatizar la identificación y la adaptación de las pautas seleccionadas. Estos métodos fueron implementados en el lenguaje de programación Python (Python Software Foundation, s. f.).

#### 2.4.1. Pauta sobre conectores complejos

El método desarrollado para la detección de conectores complejos se basa en: a) técnicas de representación de conocimiento, concretamente en sistemas de producción que tienen una estructura 'si-entonces' y en encaje de patrones; y b) en el procesamiento del lenguaje natural subsimbólico, basado en modelos del lenguaje y aprendizaje automático. Para detectar conectores complejos presentes en el texto proporcionado, se utilizan, por tanto, etiquetas proporcionadas por algoritmos de

<sup>2</sup> Una técnica (IEEE, 1990) es "un procedimiento técnico y de gestión utilizado para alcanzar un objetivo determinado".

procesamiento del lenguaje natural, así como reglas declarativas junto con unificación sintáctica.

Para adaptar los conectores complejos identificados, se ha desarrollado un método basado en el establecimiento de una tabla de equivalencias (conector complejo versus conectores simples equivalentes) y el uso de encaje de patrones.

*Sin embargo, aunque o por tanto* son algunos de los conectores complejos que pueden dificultar la comprensión de un texto para ciertos lectores. El servicio web implementado, según el método anteriormente explicado, identifica dichos conectores y los sustituye por sus equivalentes más sencillos, como *pero* o *y*. De esta manera, se facilita la lectura y se logra una mayor coherencia en el texto, mejorando su comprensión.

#### 2.4.2. Pauta sobre mayúsculas y superlativos

El servicio web realiza un análisis gramatical exhaustivo del texto para asegurarse de que las mayúsculas y los superlativos se utilicen de manera correcta y coherente. La utilización inadecuada de mayúsculas o superlativos puede cambiar el significado de una frase o generar ambigüedades, lo que afecta negativamente a la comprensión del texto. Por lo cual, el servicio verifica y corrige cuidadosamente el uso de mayúsculas en nombres propios, títulos y otros casos relevantes, así como el uso adecuado de los superlativos para expresar la máxima intensidad de una cualidad sin confundir al lector. Esta corrección contribuye a una mayor claridad y comprensión del contenido.

#### 2.4.3. Pauta sobre enumeraciones

Las enumeraciones resultan un problema, ya que varias pautas de la Norma UNE (AENOR, 2018) de lectura fácil hacen mención de ellas. Por ese motivo, se ha decidido señalar como problemáticas las enumeraciones lineales de 4 o más elementos y adaptarlas como listados con viñetas. Para identificar las enumeraciones, se ha diseñado y desarrollado un método basado en expresiones regulares, que buscan patrones típicos de enumeraciones, y relaciones de dependencia sintáctica, que confirman o desmienten si un patrón detectado es una estructura enumerativa. Estas dependencias sintácticas son etiquetas proporcionadas por algoritmos de procesamiento del lenguaje natural. Durante la adaptación se hace uso de la información recopilada en la actividad de identificación para colocar los dos puntos en el lugar adecuado, y después se segmenta la enumeración, colocando cada elemento en una viñeta del listado.

La implementación de esta pauta se evaluó con un conjunto de 141 frases extraídas del *BOE* y de artículos del periódico *Público*.

En este punto merece la pena mencionar que uno de los objetivos transversales de esta investigación es conseguir un proyecto de investigación en IA inclusivo.

Para ello, se ha incorporado en el proyecto la participación de personas con dificultades de comprensión lectora y expertos en lectura fácil, mediante el uso de cuestionarios online. Estos cuestionarios se enviaron de forma masiva el 2 de noviembre de 2023 y se cerraron el 10 de noviembre de 2023.

Por un lado, en el caso de las personas con dificultades de comprensión lectora, el cuestionario se centra en seleccionar qué opción de adaptación de enumeraciones es más fácil. El cuestionario empieza recabando información pertinente de los participantes, continúa con el test de comprensión de las adaptaciones realizadas por ATECA y finaliza con la opción de dejar un comentario voluntario. 164 personas con discapacidad cognitiva respondieron al cuestionario (el 95% eran españoles y el 5% restante provenían de Bélgica, Colombia, Francia, Rusia y Uruguay). La gran mayoría de participantes eran personas con discapacidad intelectual (79%), de las cuales un 5% también tenían alguna discapacidad física y/o enfermedad mental. En un número más reducido también respondieron al cuestionario personas con Trastornos del Espectro Autista (TEA) (3%) y Trastorno por Déficit de Atención (TDA) (2%), entre otras. Con respecto al nivel de comprensión lectora de los participantes, la mayoría respondió que comprenden 'Todo o casi todo' (43%), seguido de las personas que entendían 'Algunas cosas' (37%) y 'Poco o muy poco' (9%); el resto de participantes indicaron que su nivel de comprensión dependía del texto. Es importante resaltar que un 41% de los participantes necesitó ayuda para completar el cuestionario. Además, se puede añadir que hubo paridad de género en los participantes, hubo un 49% tanto de mujeres como de hombres. Finalmente, con respecto a la edad de los participantes, la mayoría (40%) tenían entre 31 y 45 años, seguidos de un 39% de participantes entre 17 y 30 años; el resto tenían más de 45 años.

Las respuestas a este cuestionario, junto a 31 comentarios de los participantes, han servido para tomar decisiones sobre la manera más adecuada de adaptar enumeraciones. Por otro lado, en el caso de los expertos en lectura fácil, el cuestionario enviado tenía como objetivo validar las adaptaciones realizadas por ATECA. 73 expertos en lectura fácil respondieron al cuestionario, ayudando a matizar y mejorar la implementación de esta pauta. De los 73 expertos, un 62% tenía algún tipo de discapacidad, el 51% del total tenían discapacidad intelectual y el 11% restante se divide entre discapacidad física, TEA y trastorno bipolar. Solo un 29% de los adaptadores necesitaron ayuda para completar el cuestionario. En cuestión de género, un 63% de los participantes fueron mujeres y un 34% hombres, el resto prefirió no contestar. Respecto a la edad, la gran mayoría tienen entre 31 y 45 años, un 51%, seguidos por el 31% de participantes entre 18 y 30 años. El resto tienen más de 45 años. Finalmente, mencionar que el 96% de los participantes provienen de España, y el resto de Colombia, Francia y Brasil.

#### 2.4.4. Pauta sobre palabras complejas

Con el objetivo de que ATECA contemplara de forma más adecuada la identificación de palabras complejas, se realizó un cuestionario *online* involucrando adap-

tadores de lectura fácil. Este cuestionario recoge datos relevantes sobre cómo los adaptadores realizan su trabajo de adaptación a lectura fácil, en concreto, la adaptación de palabras complejas. El cuestionario consiste en 3 secciones: 1) preguntas sobre la longitud de las palabras; 2) preguntas sobre la complejidad de las sílabas; y 3) preguntas generales sobre la pauta. El cuestionario estuvo en funcionamiento desde el 22 de noviembre de 2023 hasta el 1 de diciembre de 2023. En dicho periodo se obtuvieron 21 respuestas de adaptadores. Los resultados permitieron tomar decisiones sobre lo que se considera una palabra compleja:

- Palabras con 5 o más sílabas [excepto si se encuentran entre las 2000 palabras más frecuentes dentro del listado de la Real Academia Española (RAE) (Real Academia Española, 2023)].
- Palabras con 3 consonantes seguidas [excepto si se encuentran entre las 2000 palabras más frecuentes dentro del listado de la RAE (Real Academia Española, 2023)].
- Palabras con 4 consonantes seguidas [excepto si se encuentran entre las 1000 palabras más frecuentes dentro del listado de la RAE (Real Academia Española, 2023)].
- Palabras con patrones de sílaba CCVCC (C significa constante y V vocal), CVCC y VCC [excepto si se encuentran entre las 2000 palabras más frecuentes dentro del listado de la RAE (Real Academia Española, 2023)].

Siguiendo estos criterios, el método diseñado y desarrollado para que ATECA pueda identificar eficazmente palabras potencialmente problemáticas para la accesibilidad cognitiva se basa en técnicas de representación de conocimiento; concretamente en sistemas de producción y en encaje de patrones.

Una vez detectadas las palabras complejas, el método utilizado se basa en la búsqueda de palabras simples con el mismo significado y la posterior selección de la palabra más adecuada mediante modelos de similitud semántica. Las palabras simples se han obtenido de una lista de sinónimos, recaudados del *Diccionario de la lengua española (DLE)* (Real Academia Española, 2023), el *Tesoro de la Unesco* (Tesoro de la Unesco, s/f), *BabelNet* (BabelNet, 2023) y *SinonimosOnline.com* (s. f.).

La implementación de esta pauta se ha evaluado con un conjunto de 50 frases, generadas automáticamente con inteligencia artificial vía ChatGPT (OpenAI, 2023).

#### 2.4.5. Pauta sobre palabras indeterminadas

Palabras como *cosa*, *algo* o *asunto* pueden resultar confusas o producir ambigüedad. El método diseñado y desarrollado para la identificación de palabras indeterminadas se basa en encaje de patrones y similitud sintáctica. Esta pauta se centra en la identificación del siguiente patrón: “Verbo ‘ser’ (en cualquier forma verbal) + Palabra ‘cosa’, ‘algo’, ‘asunto’, ‘cosas’ o ‘asuntos’”. Si dicho patrón se detecta en el texto, entonces la pauta no se está cumpliendo. En ese caso, la palabra de conteni-

do indeterminado identificada se sustituye por un sinónimo adecuado al contexto. En la implementación de este método, se ha utilizado el recurso externo *BabelNet* (BabelNet, 2023) para obtener sinónimos adecuados al contexto.

La implementación de este método se ha evaluado con un conjunto de 20 frases creadas por el equipo de pruebas de ATECA.

#### 2.4.6. Pauta para evitar sinónimos

Esta pauta se implementa de la misma manera que la pauta descrita en la sección 2.4.4, es decir, utilizando el mismo sistema de extracción de sinónimos para buscar si aparecen en secciones posteriores del texto. En caso de que se encuentren varias palabras sinónimas entre sí, se marcan como problemáticas. La adaptación selecciona el primer sinónimo detectado en el texto y reemplaza los sinónimos posteriores con esa palabra. Por motivos de velocidad y desempeño, solo se tratan nombres comunes y propios. La implementación de esta pauta se ha evaluado con un conjunto de 20 frases, generadas automáticamente mediante la herramienta ChatGPT (OpenAI, 2023).

#### 2.4.7. Pauta sobre uso del imperativo

Para la corrección del uso del imperativo, se utilizan etiquetas lingüísticas (*Part-of-Speech [POS] tags* en inglés) proporcionadas por el recurso spaCy (Explosion AI, 2023) para identificar aquellos verbos que están en modo imperativo y que por el contexto se puedan confundir por la tercera persona del singular del presente de indicativo. La adaptación se realiza utilizando el patrón: “Tener” + “que” + “verbo infinitivo”. La implementación de esta pauta se ha evaluado con un conjunto de 30 frases creadas por el equipo de pruebas.

#### 2.4.8. Pauta sobre la voz pasiva

Para evitar la voz pasiva, primero se identifican las oraciones en voz pasiva empleando patrones de etiquetas de partes del discurso (*POS tagging*) y relaciones de dependencia entre palabras mediante spaCy (Explosion AI, 2023). Para refinar la precisión de esta solución, se recurre al uso de un modelo basado en la arquitectura *transformers* (modelo trf). Asimismo, se categorizaron las identificaciones en dos clases: pasivas con agente y pasivas sin agente. En lo que respecta a la adaptación, se han implementado dos enfoques distintos. En el caso de las construcciones pasivas con agente, se modifica la estructura de la oración y el verbo para convertirlas

a voz activa. Para las construcciones pasivas sin agente, dada su complejidad, se propone:

1. Utilizar la voz pasiva impersonal con *se*, ejemplificado en la transformación de *El concierto fue cancelado* a *Se canceló el concierto*.
2. Emplear un sujeto indefinido como *alguien*, transformando *El concierto fue cancelado* a *Alguien canceló el concierto*.

Para la evaluación de la pauta, se hizo uso de 70 frases redactadas por el equipo de pruebas de ATECA o generadas automáticamente mediante ChatGPT (OpenAI, 2023).

## 2.5. Integración de pautas en el sistema piloto

Todos los métodos desarrollados se integraron en la interfaz de usuario descrito en la sección 2.3. Como resultado de dicha integración, se ha obtenido la aplicación web piloto denominada "ATECA: Adaptación de TExtos en CAstellano". Esta aplicación se ha desplegado en un servidor web de la UPM el día 22 de diciembre de 2023. A partir de ese momento, ATECA está disponible para que los adaptadores de lectura fácil o interesados en la accesibilidad cognitiva puedan utilizar el piloto. La Figura 8 muestra una captura de pantalla de ATECA cuando el usuario está usando la funcionalidad de identificar problemáticas con respecto a la Norma UNE de lectura fácil (AENOR, 2018).

La Figura 9 muestra la interfaz de la funcionalidad de proponer sugerencias de adaptación. Cada adaptación del texto es resaltada según si la adaptación tiene que ser manual (amarillo), la adaptación sugerida sigue la Norma UNE (AENOR, 2018) (verde) o hay varias opciones de adaptación (morado).



FIGURA 9. ATECA: Sugerencias de adaptación



### 3. Resultados

El sistema piloto fue desplegado en la dirección <https://ateca.linkeddata.es> el 22 de diciembre de 2023 para el uso por parte de adaptadores de texto. Adicionalmente, el sistema fue conectado a la plataforma Google Analytics, permitiendo la recogida de estadísticas de uso y de registro entre el 8 de enero y el 26 de junio de 2024.

Para llevar a cabo la difusión del piloto, se han realizado diversas actividades en estrecha colaboración entre el equipo de trabajo de la UPM y el equipo de trabajo del CEACOG. En concreto, se impartió una charla de concienciación a profesionales de la lectura fácil para: 1) explicar la investigación y el desarrollo que se está llevando a cabo en la UPM para construir un sistema piloto de ayuda en el proceso de adaptación a lectura fácil; y 2) mencionar los beneficios de este sistema piloto, mostrando su alcance actual. La charla tuvo lugar de forma remota el día 20 de noviembre de 2023 utilizando la plataforma Zoom. La charla duró 70 minutos y tuvo un total de 45 asistentes (sin contar a la ponente).

Además, se creó una sección de difusión en la página web del CEACOG. En dicha sección se presenta una explicación de la investigación realizada y un enlace a ATECA, para que pueda usarse por profesionales del ámbito de la discapacidad.

También se elaboró una nota de prensa para explicar en detalle el proyecto de investigación y su principal resultado, ATECA.

Las siguientes secciones muestran las estadísticas de registro (sección 3.1) y uso (sección 3.2), respectivamente. Además, la sección 3.3 presenta ejemplos de uso de ATECA.

### 3.1. Estadísticas de registro

Entre las fechas del 8 de enero al 26 de junio de 2024, hubo un total de 1020 registros, que presentaron los siguientes perfiles:

- *País de residencia*: España (88%), Argentina (5%), Colombia (2.35%), Chile (1.27%) y México (1.27%); también presentes Brasil, Costa Rica, El Salvador, EE. UU. y Suiza, con menos del 1%.
- *Comunidades autónomas* (cuando aplica): Andalucía (17%), Madrid (13%), Valencia (11%), Galicia (8%), Navarra (8%), Cataluña (7%), Castilla y León (6%) y País Vasco (6%); también presentes Aragón, Asturias, La Rioja, Canarias, Galicia, Castilla-La Mancha, Baleares, Cantabria y Ceuta con menos del 5%.
- *Uso de ATECA*: profesional (68%), personal (26%) y otro (6%).
- *Experiencia con lectura fácil*: el 34% tiene experiencia previa frente al 66% que no.
- *Miembros de Plena inclusión*: el 14% pertenece a Plena frente al 86% que no.

### 3.2. Estadísticas de uso

Como el uso del sistema se encuentra dividido en dos operaciones, identificación y adaptación, se han registrado tanto el número total de llamadas a cada operación, así como el número de veces que las pautas disponibles han sido seleccionadas en una operación. Estos datos pueden ser consultados en la Tabla 1. Cabe destacar que una operación puede, y suele, tener seleccionadas simultáneamente más de una pauta.

La Tabla 1 recoge la actividad del uso de ATECA, según si han utilizado Identificación (1734) o Adaptación (2869) y qué pautas han querido corregir.

### 3.3. Uso de ATECA

Como se ha mencionado en las secciones 2.3 y 2.6, ATECA permite a los adaptadores de lectura fácil: a) identificar y analizar los problemas presentes en el texto original que puedan dificultar su comprensión y que no cumplan con la Norma UNE de lectura fácil (AENOR, 2018); y b) proponer sugerencias de adaptación para hacer el texto original más accesible y comprensible, y acorde a las pautas de lectura fácil.

TABLA 1. Estadísticas de uso

Usos	Identificación	Adaptación
Pauta Conectores	1336	2091
Pauta Enumeraciones	1273	1949
Pauta Evitar sinónimos	1253	1954
Pauta Uso del imperativo	1300	2000
Pauta Mayúsculas	1262	1846
Pauta Palabras complejas	1282	2031
Pauta Palabras indeterminadas	1328	1997
Pauta Superlativos	1365	2008
Pauta Voz pasiva	1363	2149
Totales	1734	2869

La Figura 10 muestra el uso de la opción “Identificar Incumplimientos de la Norma UNE”, donde se identifican los casos problemáticos del uso de mayúsculas (“No se deben escribir palabras ni frases con todas sus letras en mayúsculas, excepto cuando se trate de siglas”) y superlativos (“Se deben evitar los superlativos. Es recomendable añadir el adverbio muy al adjetivo o al adverbio”) en el texto introducido. ATECA presenta un breve informe en el que se muestran las pautas que se cumplen en el texto original, marcadas en verde, y las pautas que no se cumplen, marcadas en rojo.

Seleccionando la opción “Obtener sugerencias para adaptar el texto”, ATECA devuelve el texto introducido con una sugerencia de cómo corregir las pautas seleccionadas.

La Figura 11 ofrece un ejemplo de adaptación de imperativos (“Se debería utilizar el imperativo solo en contextos claros para evitar que se confunda con la tercera persona del singular del presente de indicativo”), donde ATECA sugiere el cambio de *Mirad* en el texto original a *Tenéis que mirar* en el texto adaptado, resaltado en verde.

El ejemplo de la Figura 12 expone un caso donde se quiere obtener sugerencias de adaptación de frases en voz pasiva y de frases que contengan enumeraciones de elementos en un texto original introducido en la caja de texto de la izquierda. El texto adaptado por ATECA, que aparece en la caja de texto de la derecha, presenta una propuesta de adaptación de la enumeración presente en el texto original (en verde) y una sugerencia de cambio para la frase en voz pasiva (en morado).

FIGURA 10. Identificación de mayúsculas y superlativos

The screenshot displays the ATECA web application interface. At the top, the logo 'ATECA Piloto' is visible, along with the text 'Adaptación de Textos en Castellano' and a 'Inicio' button. A 'Cerrar Sesión' button is located in the top right corner. The main content area features two primary buttons: 'Identificar incumplimientos de la Norma UNE' and 'Obtener sugerencias para adaptar el texto'. Below these, a section titled 'Selecciona las pautas con las que quieres trabajar:' includes a 'Seleccionar todo' button and a list of checkboxes for various adaptation rules. The 'Mayúsculas' checkbox is checked. A text editor window shows a sample text about Santiago Scarlati, with a red box highlighting the word 'muchísimas'. Below the text editor, a button reads 'Hay un informe a continuación sobre las pautas de lectura fácil'. At the bottom of the interface, there are buttons for 'Otro texto' and 'Identificar'. A notification box indicates a violation of the 'Superlativos' rule, stating that the UNE recommendation advises avoiding superlatives and suggesting the use of adverbs instead. It lists 'rapidísimo' and 'muchísimas posibilidades' as elements that do not follow the norm. A second notification box shows a green checkmark for 'Mayúsculas'. The footer contains logos of various institutions and a contact email: 'easy2read@delicias.dia.fi.upm.es'.

#### 4. Conclusiones y líneas futuras

La colaboración entre la UPM y el CEACOG ha permitido el desarrollo de una aplicación web piloto llamada ATECA. ATECA es un sistema de apoyo basado en IA que se puede incorporar como ayuda en el proceso global de adaptación de materiales a las pautas indicadas en la metodología de lectura fácil. Este sistema de apoyo permite realizar las funciones de identificación y corrección de textos a lectura fácil, lo que puede facilitar el proceso de adaptación de textos a los adaptadores.

Actualmente, ATECA realiza la identificación y la sugerencia de correcciones para un conjunto de 9 pautas de la metodología lectura fácil. Estas pautas cubren aspectos fundamentales del lenguaje, tales como el vocabulario, la estructura gramatical de frases y el uso correcto de conectores o marcadores discursivos. La

FIGURA 11. Adaptación de imperativos

Identificar incumplimientos de la Norma UNE

Obtener sugerencias para adaptar el texto

Selecciona las pautas con las que quieres trabajar:

Seleccionar todo Deseleccionar todo

Vocabulario y expresiones  Frases y palabras  Ortografía

Palabras complejas  Voz pasiva  Mayúsculas

Superlativos  Conectores complejos  Enumeraciones

Palabras indeterminadas  Imperativos

Evitar síndromos

Sana Serif Normal B I U A X

Mirad cómo se golpea el pecho, y cómo el otro, suspirando, apoya su mejilla en la palma de la mano. Padre y suegro son del mal de Francis: saben que su vida es grossera y viciosa, y de ahí proviene el dolor que les aflige.

Tenéis que mirar cómo se golpea el pecho, y cómo el otro, suspirando, apoya su mejilla en la palma de la mano. Padre y suegro son del mal de Francis: saben que su vida es grossera y viciosa, y de ahí proviene el dolor que les aflige.

Eliminar colores

Este texto podría necesitar una adaptación manual

Tu texto se puede adaptar usando la sugerencia propuesta

Hay más de una sugerencia para adaptar tu texto

Otro texto Sugerir

FIGURA 12. Adaptación de voz pasiva y enumeraciones

Identificar incumplimientos de la Norma UNE

Obtener sugerencias para adaptar el texto

Selecciona las pautas con las que quieres trabajar:

Seleccionar todo Deseleccionar todo

Vocabulario y expresiones  Frases y palabras  Ortografía

Palabras complejas  Voz pasiva  Mayúsculas

Superlativos  Conectores complejos  Enumeraciones

Palabras indeterminadas  Imperativos

Evitar síndromos

Sana Serif Normal B I U A X

Tanto la Fundación Lanteri Bataui como Teuibinsa presentan programas muy bien establecidos en lo referente al trabajo con PCDI en proceso de deterioro por envejecimiento en el entorno laboral. Ambas entidades llevan trabajando en este ámbito ya varios años, cuentan con herramientas propias y específicas de valoración del deterioro laboral debido al envejecimiento, con personal concreto dedicado a este trabajo y con un número significativo de PCDI que se han beneficiado de su participación y con sistemas de evaluación de los resultados obtenidos. De esta manera y tal y como se expone más adelante, ambas entidades presentan programas que podrían ser considerados como Buenas Prácticas en este campo.

Tanto la Fundación Lanteri Bataui como Teuibinsa presentan programas muy bien establecidos en lo referente al trabajo con PCDI en proceso de deterioro por envejecimiento en el entorno laboral. Ambas entidades llevan trabajando en este ámbito ya varios años, cuentan con herramientas propias y específicas de valoración del deterioro laboral debido al envejecimiento con personal concreto dedicado a este trabajo y con un número significativo de PCDI que se han beneficiado de su participación con sistemas de evaluación de los resultados obtenidos. De esta manera y tal y como se expone más adelante, ambas entidades presentan programas que podrían considerarse como Buenas Prácticas en este campo.

Eliminar colores

Este texto podría necesitar una adaptación manual

Tu texto se puede adaptar usando la sugerencia propuesta

Hay más de una sugerencia para adaptar tu texto

Otro texto Sugerir

capacidad de ATECA para ofrecer sugerencias precisas y personalizadas convierte a esta herramienta en un apoyo valioso para los adaptadores, quienes pueden ver acelerado su trabajo diario.

Esta aplicación web piloto fue lanzada el 22 de diciembre de 2023. Desde dicho lanzamiento, ATECA ha sido utilizada por más de 1000 usuarios con diferentes perfiles y experiencias en la adaptación de textos. Este nivel de adopción demuestra que existe una demanda significativa de herramientas tecnológicas que mejoren y optimicen el proceso de adaptación de documentos a lectura fácil. Los datos de uso, recogidos a través de Google Analytics, confirman que las funcionalidades de identificación y sugerencia de adaptaciones han sido ampliamente empleadas, lo que subraya la utilidad práctica de ATECA en entornos profesionales y personales.

Desde el lanzamiento de este piloto, se han realizado diferentes actualizaciones para mejorar el funcionamiento de ATECA. Algunas de esas mejoras se han realizado a propuesta de los usuarios de ATECA, como, por ejemplo, la inclusión de un botón para desmarcar las pautas seleccionadas o la incorporación de un nuevo color para indicar una problemática en el texto para la que ATECA no ha podido sugerir una adaptación adecuada.

Sin embargo, aún existen áreas de mejora que deben abordarse en futuros desarrollos. De cara al futuro, las líneas de investigación y desarrollo deben centrarse en mejorar la precisión de las adaptaciones mediante el análisis y la aplicación de nuevas técnicas que aúnen procesamiento de lenguaje natural y aprendizaje automático. Se contempla también la expansión de las funcionalidades del sistema, incorporando pautas adicionales y permitiendo una mayor flexibilidad para ajustarse a las necesidades específicas de diferentes usuarios y tipos de texto. Estas mejoras se basarán en la realización de una investigación multidisciplinar que aúne lingüística, cognición y tecnología, y que involucre a profesionales de distintas áreas para una mayor eficacia a la hora de identificar y corregir problemáticas de lectura fácil en textos en castellano.

Asimismo, es fundamental continuar con la evaluación de ATECA mediante pruebas en entornos reales y con una amplia gama de usuarios, lo que permitirá identificar oportunidades de mejora y validar la eficacia de las nuevas funcionalidades.

En resumen, ATECA ha dado un paso importante hacia la mejora de la accesibilidad cognitiva en la adaptación de textos en español. Aunque aún se encuentra en una fase de desarrollo, los resultados obtenidos son alentadores y muestran que herramientas basadas en inteligencia artificial, como ATECA, pueden desempeñar un papel crucial en la democratización del acceso a la información.

## 5. Referencias bibliográficas

- AENOR, ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN. (2018). *Lectura Fácil. Pautas y recomendaciones para la elaboración de documentos* (UNE: 153101 EX).<https://www.une.org/encuentra-tu-norma/busca-tu-norma/norma/?c=N0060036>
- BABELNET. (2023). Entry Title. <https://babelnet.org>

- CARGO, M. y MERCER, S. L. (2008). The value and challenges of participatory research: Strengthening its practice. *Annu. Rev. Public Health*, 29, 325-350.
- EXPLOSION AI. (2023). spaCy. <https://spacy.io/>
- IEEE STANDARD GLOSSARY OF SOFTWARE ENGINEERING TERMINOLOGY (1990). IEEE Std. 610.12-1990 (Revision and redesignation of IEEE Std. 792-1983).
- IEEE STANDARD GLOSSARY OF DATA MANAGEMENT TERMINOLOGY (1990b). IEEE Std. 610.5-1990.
- OPENAI. (2023). CHATGPT (versión del 15 de julio) [Modelo de lenguaje de gran tamaño]. <https://chat.openai.com/chat>
- PLENA INCLUSIÓN. (s. f.). *¿Qué es la lectura fácil?* <https://www.plenainclusion.org/discapacidad-intelectual/recurso/lectura-facil/>
- PYTHON SOFTWARE FOUNDATION. (s. f.). Python. <https://www.python.org/>
- REAL ACADEMIA ESPAÑOLA. (2023). *Diccionario de la lengua española* (23.ª ed., versión 23.7 en línea). <https://dle.rae.es> [Fecha de consulta: 10/2023].
- RUSSELL, S. y NORVIG, P. (2021). *Artificial Intelligence: A Modern approach, global edition*. Pearson Higher Ed.
- SinonimosOnline.com. (s. f.). SinonimosOnline.com. Recuperado de <https://www.sinonimosonline.com/>
- SUÁREZ-FIGUEROA, M. C., DIAB, I., RUCKHAUS, E. y CANO, I. (2024). First steps in the development of a support application for easy-to-read adaptation. *Universal Access in the Information Society*, 23(1), 365-377. Publicado *online*: 18 November 2022. <https://doi.org/10.1007/s10209-022-00946-z>
- TESAURO DE LA UNESCO (s/f). <https://vocabularies.unesco.org/browser/thesaurus/es/>

## Anexo

- Políticas de protección de datos que se utilizan en ATECA: <https://oeg.fi.upm.es/protecciondatos.html>
- Cuestionario sobre enumeraciones respondido por personas con discapacidad cognitiva: <https://forms.gle/ueHfTrpk2bWATFKd8>
- Cuestionario sobre enumeraciones respondido por adaptadores: <https://forms.gle/eGCVPNJs4RVobUne8>
- Cuestionario de palabras complejas respondido por adaptadores: <https://forms.gle/t7nc5caqPibR8dcq8>
- Diccionario de la lengua española: <https://dle.rae.es/>
- Tesauruso de la UNESCO: <https://vocabularies.unesco.org/browser/thesaurus/es/>
- BabelNet: <https://babelnet.org/about>
- sinonimosonline: <https://www.sinonimosonline.com/>
- spaCy: <https://spacy.io/>
- Lista total de frecuencias del español: <https://corpus.rae.es/lfrecuencias.html>
- Página de difusión de ATECA en la página web del CEACOG: <https://www.ceacog.es/que-hacemos/aplicacion-para-adaptar-textos-a-lectura-facil/>
- Nota de prensa sobre ATECA: <https://www.ceacog.es/2024/01/31/derechos-sociales-desarrolla-ateca-la-inteligencia-artificial-para-adaptar-textos-a-lectura-facil/>
- ATECA: <https://ateca.linkeddata.es>
- Google Analytics: <https://analytics.google.com>