



**UNIVERSIDAD POLITÉCNICA SALESIANA**  
**SEDE CUENCA**  
**CARRERA DE COMPUTACIÓN**

DISEÑO Y DESARROLLO DE UN MÓDULO DE ADAPTACIÓN A LECTURA FÁCIL  
PARA LA HERRAMIENTA DE ADAPTACIÓN DE OBJETOS DE APRENDIZAJE DEL  
PROYECTO EDUTECH

Trabajo de titulación previo a la obtención del  
título de Ingeniero en Ciencias de la Computación

AUTORES: EVELYN SOLEDAD PINTADO CHACHO

ADRIÁN FERNANDO VISCAINO CONCE

TUTORA: ING. PAOLA CRISTINA INGAVÉLEZ GUERRA, MSc.

Cuenca - Ecuador

2022

## **CERTIFICADO DE RESPONSABILIDAD Y AUTORÍA DEL TRABAJO DE TITULACIÓN**

Nosotros, Evelyn Soledad Pintado Chacho con documento de identificación N° 1720171337 y Adrián Fernando Viscaino Conce con documento de identificación N° 0106171325; manifestamos que:

Somos los autores y responsables del presente trabajo; y, autorizamos a que sin fines de lucro la Universidad Politécnica Salesiana pueda usar, difundir, reproducir o publicar de manera total o parcial el presente trabajo de titulación.

Cuenca, 07 de marzo del 2022.

Atentamente,

---

Evelyn Soledad Pintado Chacho

1720171337

---

Adrián Fernando Viscaino Conce

0106171325

## **CERTIFICADO DE CESIÓN DE DERECHOS DE AUTOR DEL TRABAJO DE TITULACIÓN A LA UNIVERSIDAD POLITÉCNICA SALESIANA**

Nosotros, Evelyn Soledad Pintado Chacho con documento de identificación N° 1720171337 y Adrián Fernando Viscaino Conce con documento de identificación N° 0106171325, expresamos nuestra voluntad y por medio del presente documento cedemos a la Universidad Politécnica Salesiana la titularidad sobre los derechos patrimoniales en virtud de que somos autores del Artículo Académico: “Diseño y desarrollo de un módulo de adaptación a lectura fácil para la herramienta de adaptación de objetos de aprendizaje del proyecto EduTech”, el cual ha sido desarrollado para optar por el título de: Ingeniero en Ciencias de la Computación, en la Universidad Politécnica Salesiana, quedando la Universidad facultada para ejercer plenamente los derechos cedidos anteriormente.

En concordancia con lo manifestado, suscribimos este documento en el momento que hacemos la entrega del trabajo final en formato digital a la Biblioteca de la Universidad Politécnica Salesiana.

Cuenca, 07 de marzo del 2022.

Atentamente,

---

Evelyn Soledad Pintado Chacho

1720171337

---

Adrián Fernando Viscaino Conce

0106171325

## **CERTIFICADO DE DIRECCIÓN DEL TRABAJO DE TITULACIÓN**

Yo, Paola Cristina Ingavélez Guerra con documento de identificación N° 1712214616, docente de la Universidad Politécnica Salesiana, declaro que bajo mi tutoría fue desarrollado el trabajo de titulación: DISEÑO Y DESARROLLO DE UN MÓDULO DE ADAPTACIÓN A LECTURA FÁCIL PARA LA HERRAMIENTA DE ADAPTACIÓN DE OBJETOS DE APRENDIZAJE DEL PROYECTO EDUTECH, realizado por Evelyn Soledad Pintado Chacho con documento de identificación N° 1720171337 y por Adrián Fernando Viscaino Conce con documento de identificación N° 0106171325, obteniendo como resultado final el trabajo de titulación bajo la opción Artículo Académico, que cumple con todos los requisitos estipulados por la Universidad Politécnica Salesiana.

Cuenca, 07 de marzo del 2022.

Atentamente,

---

Ing. Paola Cristina Ingavelez Guerra, MSc.

1712214616

## **DEDICATORIA Y AGRADECIMIENTO**

*Agradezco de todo corazón a mi mami Soledad que ha sido el pilar fundamental y la persona que me ha dado fuerzas, apoyo incondicional y sobre todo amor y paciencia y que sin su apoyo no sería nada de lo que soy ahora, gracias por confiar en mí cuando ni yo misma lo hacía y definitivamente este es un esfuerzo a la par por lo que el título es de las dos, espero seguir haciéndola sentir orgullosa con mi trabajo y definitivamente todo su esfuerzo será recompensado. También agradezco a mi abuelita por haberme criado con los valores que eh tenido siempre presente y a mi hermana y mi sobrina, quienes me han enseñado a madurar y crecer como persona y a mi familia que de una u otra forma han logrado que esto sea posible.*

*También estoy muy agradecida por las personas más geniales que eh conocido en esta etapa de mi vida con los que recorrimos todo este proceso y les doy gracias a Cris, Edwin, Pablo y Adrián por permitirme formar parte de su vida, por siempre los llevaré en mi corazón.*

***Evelyn Soledad Pintado Chacho***

*Primero agradezco a Dios por darme la sabiduría suficiente para culminar con éxito esta etapa de mi vida, la fuerza para enfrentar cada obstáculo que se me ha impuesto y las ganas para salir adelante. Agradezco a mis padres que me han enseñado el valor de obtener las cosas por uno mismo con esfuerzo y trabajo y que solo dependo de mí mismo para conseguir lo que quiero. A mi hermana por enseñarme con el ejemplo todos los aspectos de la vida especialmente en el ámbito profesional. A mis maestros que me han enseñado lo necesario para poder triunfar brindándome las bases necesarias e inculcando en mí la investigación. Finalmente doy gracias a mis amigos y compañeros que me han apoyado de varias formas a cumplir mis objetivos y sueños.*

***Adrian Fernando Viscaino Conce***

## **Resumen**

La dificultad de la comprensión de texto para personas de grupos específicos supone un reto en las nuevas prácticas de enseñanza por lo que la lectura fácil nace como una opción de accesibilidad para estas. Lo que recomiendan es que se utilice o se filtre la lectura según el perfil de los usuarios, lo que facilitaría notablemente la lectura, pero teniendo en cuenta que no todos los perfiles están desarrollados en su totalidad para el público en general, este trabajo realiza una investigación para ayudar a mitigar la brecha entre usuarios generalizando su uso para un público más extenso. En base a la investigación se encontraron pautas que siguen varios programas creando de esta forma un estándar para su uso. En base a estas pautas se buscó crear una estructura de redes neuronales que brinden una arquitectura de como formar oraciones sencillas en base a OAs (Objetos de Aprendizajes) planteados siendo su finalidad ser un módulo que se implementa posteriormente en el proyecto EDUTECH. El plan de experimentación para la verificación de los análisis de los resultados se basa en fórmulas de legibilidad y comprensibilidad que serán validadas con el módulo. Finalmente cuenta con una aplicación básica para pruebas y un mejor análisis de los resultados facilitando su uso.

**Palabras claves:** Lectura fácil, PLN, procesamiento del lenguaje natural, redes neuronales, standford nlp, stanza nlp, svm linear.

## **Abstract**

The difficulty of text comprehension for people from specific groups is a challenge in new teaching practices, so easy reading is born as an accessibility option for them. What they recommend is that reading be used or filtered according to the users' profile, which would undoubtedly facilitate reading, but bearing in mind that not all profiles are fully developed for the public, this work research to help bridge the gap between users by generalizing its use to a broader audience. Based on the investigation, guidelines will be found that several programs follow, thus creating a standard for their use. Based on these guidelines, it was sought to create a structure of neural networks that provide an architecture of how to form simple sentences based on OAs (Learning Objects) proposed, its purpose being a module that will be implemented later in the EDUTECH project. The experimentation plan for the verification of the analysis of the results, it is based on readability and comprehension formulas that will be validated with the module. Finally, it has a basic application for tests and a better analysis of the results, facilitating its use.

**Keywords:** Easy Reading, PLN, natural language processing, neural networks, standford nlp, stanza nlp, svm linear.

## Índice de Contenido

<b>1. Introducción</b> .....	1
<b>2. Antecedentes</b> .....	1
<b>2.1. Accesibilidad para personas con discapacidad intelectual en el Ecuador</b> .....	1
<b>2.2. Impacto de la Lectura Fácil</b> .....	2
<b>3. Trabajos Relacionados</b> .....	4
<b>3.1. Accesibilidad Web</b> .....	4
<b>3.2. Técnicas de Evaluación de textos a Lectura Fácil</b> .....	4
<b>3.3. Normas de Lectura Fácil</b> .....	8
<b>3.3.1. UNE 153101:2018 EX – Lectura Fácil: Pautas y Recomendaciones para la elaboración de documentos</b> .....	8
<b>3.3.2. Manual de Lectura Fácil y Formatos Accesibles del Ecuador</b> .....	9
<b>4. Metodología</b> .....	10
<b>4.1. Conceptos previos y recursos utilizados</b> .....	10
<b>4.2. Diseño y desarrollo del módulo</b> .....	13
<b>4.3. Diagrama del módulo de adaptación</b> .....	19
<b>4.4. Despliegue en entornos definidos</b> .....	21
<b>5. Resultados y evaluaciones</b> .....	21
<b>6. Conclusiones y Trabajo Futuro</b> .....	30
<b>7. Referencias</b> .....	30



## **1. INTRODUCCIÓN**

La constante evolución de las tecnologías de información ha llevado a que internet se convierta en una de las fuentes de información más grandes conocidas por lo que es fácil encontrar información en distintos formatos de audio, video, textuales. El contenido normalmente presentado contiene textos, que a menudo es comprensible para cualquier lector pero que se vuelve complicado para ciertos tipos de personas, esto ha llevado a la sociedad a preguntarse el nivel de compromiso que se tiene a favor de las personas con discapacidad y a pensar de manera específica en las dificultades que esto conlleva.

La lectura fácil se define por primera vez como una “adaptación lingüística del texto para facilitar la lectura del texto y su comprensión” en la publicación de “Directrices para materiales de lectura fácil” (Nomura et al., 2010) publicada en 1997; aunque el término fue utilizado por primera vez en el libro de Lectura Fácil publicado en Suecia en 1968. Luego pasaron varios años y varios países trabajaron en diferentes pautas y recomendaciones sobre la lectura fácil creando de esta forma la red internacional Easy-to-Read en 2005.

Basándose en este pensamiento, surge en España la norma UNE 153101 en la que se encuentran normas y pautas generalizadas para un correcto manejo de la información y de esta forma fomentar la inclusión en la adaptación de textos; luego aparecieron proyectos de este tipo en América empezando desde México y luego con la CONADIS en Ecuador, que creó el “Manual de Lectura Fácil y Formatos Accesibles”.

Como una solución alternativa al proceso de adaptación de contenidos EduTech de Erasmus+ desarrolla herramientas que faciliten la adaptación de objetos de aprendizaje con distintas técnicas, entre ellas Lectura Fácil, requiriendo un módulo que lo realice de manera automática y que presente resultados óptimos.

## **2. ANTECEDENTES**

## **2.1. Accesibilidad para personas con discapacidad intelectual en Ecuador**

La discapacidad cognitiva o intelectual en las personas provoca un déficit en el grado de comprensión de lectura de textos, esto debido a que las estructuras son demasiado complejas, es difícil entender analogías, o el uso de expresiones que no son escuchadas diariamente, es por ello por lo que organizaciones nacionales como la CONADIS e internacionales como la Fundación Visibilia, Lectura Fácil MX, DISCAPNET, etc., a través de estudios han ido desarrollando pautas, normas y manuales para adaptar los contenidos de información a un formato que sea legible y de fácil entendimiento, esta solución ha sido determinada como Lectura Fácil. Este tipo de accesibilidad se entiende como el apoyo de una herramienta tecnológica lo que básicamente apoyará a que las personas, en todos los ámbitos pueden desenvolverse mejor. (Tuteladas et al., 2009)

La CONADIS en Ecuador con la finalidad de brindar igualdad de oportunidades considera que el acceso a la educación, información y comunicación conlleva a una verdadera inclusión por lo que constantemente buscan formas para que las personas con discapacidad estén en las mismas condiciones que cualquiera; de esta forma nace su manual para Lectura Fácil enfocándose en las personas con discapacidad intelectual o cualquier persona que tenga dificultad para comprender información, este manual hace énfasis en la importancia de la Lectura Fácil y el derecho para el acceso a la información y como este derecho se ha visto vulnerado, por lo que sus normas ayudan de diferente forma a identificar maneras en las que estas estrategias puedan resultar más adecuadas para este grupo en específico. (Tuteladas et al., 2009)

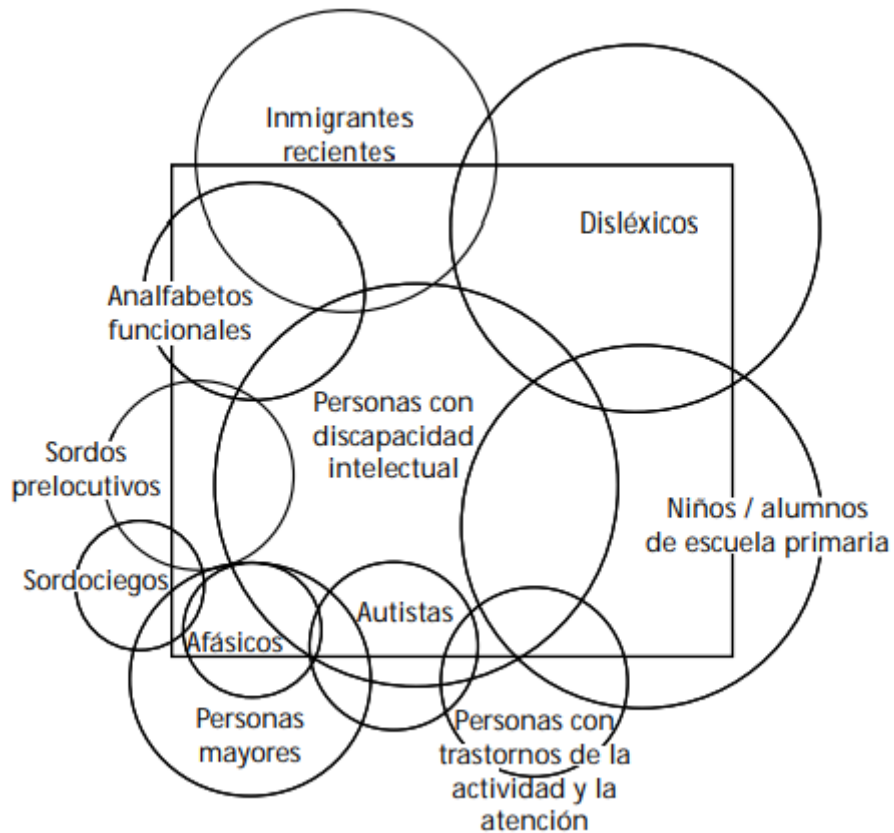
## **2.2. Impacto de la Lectura Fácil en la educación**

Es importante señalar que la idea por el acceso a la información para todos nació en los años 80 pero en la actualidad ha visto su auge gracias a las diferentes herramientas tecnológicas con las que se cuenta por lo tanto el desarrollo de la idea ha evolucionado aportando de esta forma varios usos y alcances que pueden lograrse.

El más importante de todos es el acceso a la educación ya que se ha vuelto esencial en el desarrollo pedagógico y como no está dirigida únicamente a un grupo en específico es fácilmente adaptable. De los mayores avances en esta área consideramos el trabajo que han realizado diversos países en conjunto con la Red Iberoamericana Responsable de Planes y Políticas de Lectura, su objetivo es realizar censos para analizar el nivel escolar y los planes de lectura de cada país, por ejemplo, en 2019 de un total de 20 países 12 de estos cuentan con planes de lectura sin embargo 4 de estos no cuentan con planes de lectura y otros 4 no están activos; y de estos 12 países solo 4 están dentro del marco normativo nacional educativo. (Ministerio de Educación, 2021)

Ecuador en 2017 pone en marcha el Plan Nacional de Promoción del Libro y la Lectura “José de la Cuadra” (Presidencia de la República del Ecuador, 2017) con el propósito de desarrollar las capacidades que ayudan a la interpretación de los textos, porque consideran a la lectura como una práctica individual y colectiva para un buen desarrollo social. En conjunto con la Agencia Española de Cooperación para el Desarrollo el 18 de noviembre de 2018, en Guayaquil, se capacitaron a trabajadores de las entidades públicas en lectura fácil como forma de propuesta para nuevas estrategias y metodologías para la inclusión en la sociedad.

En base a los datos anteriores, se constata que es muy poco el apoyo que se le da a los diversos grupos con diferentes perfiles los cuales presentan limitaciones y también son aptos para este tipo de lectura, ahora estos son los diversos grupos a los cuales está dirigida este tipo de lectura.



**Figura 1.** Usuarios de lectura fácil. Fuente: Asociación de Lectura Fácil, a partir de «Directrices para materiales de lectura fácil» de la IFLA

### **3. TRABAJOS RELACIONADOS**

#### **3.1 Accesibilidad Web**

Según la OMS el 15% de la población mundial cuenta con algún tipo de discapacidad y de estos son los países en vías de desarrollo los que tienen un menor índice tanto social como económico por lo que las oportunidades son pocas o mínimas, entre los obstáculos que enfrentan están el acceso a la salud y a la educación, lo que dificulta notablemente el acceso a la comunicación. (Harari & Amadeo, 2020). Desde este punto se entiende que el acceso a la información no está adaptado para todos y eso es lo que se quiere conseguir con la accesibilidad entonces se entiende

como la forma de promover el acceso a la información utilizando herramientas tecnológicas regidos bajo los lineamientos de la Web Accesability Initiative.

La lectura fácil pretende dar una solución de accesibilidad universal para favorecer a todas las personas posibles, por ejemplo, en las escuelas independientemente de los casos se intenta que la información correcta llegue a todos, de todas las maneras posibles ya que esta es la apuesta a la lectura fácil, la inclusión. De este modo el personal docente busca la forma para incluir a este tipo de personas y en su búsqueda encuentran herramientas en la web que permite el apoyo a este tipo de lectura.

En la web, las páginas de apoyo para lectura fácil cuentan con una comunidad que brinda la ayuda necesaria por medio de contactos y emails; con diferencia a estos grupos nuestro trabajo trata de brindar un apoyo inmediato en la herramienta web, evitando el proceso manual de las páginas de apoyo y realizándolo de forma automática.

### **3.2. Técnicas de Evaluación de textos a Lectura Fácil**

Para la verificación de resultados, se ha realizado un análisis lingüístico utilizando las medidas de legibilidad, comprensibilidad, perspicuidad y lecturabilidad en base a la terminología de la lengua española.

- **Legibilidad:** Según la RAE (Real Academia de la Lengua Española) lo define como “una cualidad de lo que es legible... por lo escrito o impreso comprendiendo la significación de los caracteres empleados” se lo considera entonces como un atributo del texto favoreciendo la comprensión lectora.(Ríos, 2017) Existen dos tipos de legibilidad la tipográfica y la lingüística. La tipográfica habla de la parte estética, la longitud, espaciado, justificación, entre otras; mientras que la parte lingüística funciona con palabras y frases personales.

Existen fórmulas y test para el análisis de legibilidad en el texto, al principio se contaban cuantas veces y en que rango estaban las palabras en el texto por lo que es mejor un número de palabras reducido y tener un menor número de sílabas. En 1943 Rudolf Flesch presento

la primera fórmula para la facilidad de lectura que tomaba en cuenta el grado de comprensión del texto y el factor humano y de esta forma mide la legibilidad de la lectura en adultos. Mientras que en 1953 Taylor crea el test de Cloze el que realiza un conteo de sílabas aplicando fórmulas para la lengua inglesa. (Campus, 2020)

José Antonio García López en 2001 partiendo del proyecto de Flesch, crea un método para la legibilidad basada en las edades. Por lo tanto, el toma como muestra 100 palabras procurando que sea desde el inicio del párrafo, luego se cuentan el total de sílabas y el promedio de sílabas por palabras, este proceso se denomina SIL (Campus, 2020). Luego se cuenta el número de frases y de estas el número total de palabras y a este procedimiento se lo denomina como PAL. Ya con estos valores, los aplicamos a la fórmula donde E = Edad, PAL= Conteo de Frases y SIL = Conteo de Sílabas, de la siguiente forma:

$$E = -7,1395 + 0,2495 PAL + 6,4763 SIL$$

Esta fórmula devuelve la edad de nivel de comprensibilidad del texto, ahora utilizando la fórmula de LEG obtenemos la fórmula general.

$$LEG = 93,0941 - 7,3289 E$$

$$EDAD = 12,3578 - 0,1162 LEG$$

Estas fórmulas fueron utilizadas para medir la legibilidad en las páginas web de lectores en general.

- **Comprensibilidad:** Según la RAE lo define como “calidad de comprensible”, entonces podemos tratar a la comprensibilidad como el nivel de esfuerzo que requieren las personas al entender un texto. Comúnmente *comprensible* y *legible* son utilizados para lo mismo por lo que tienen un significado parecido, pero es definido en este contexto como el acto de comunicación de los escritores para que los lectores con la finalidad de un mismo entendimiento.(Adelbery & Razek, 1984)

Spaulding en 1951, desarrolló el primer procedimiento para la evaluación de la comprensibilidad de los textos en castellano, luego de esto fue aumentando la calidad de su evaluación incluyendo en 1956 el factor promedio de la longitud y la densidad de la oración y en 1972 simplificando el nivel de dificultad de los textos para lectores con bajos niveles de conocimiento. Fray, en 1968, desarrolló un gráfico para la comprensibilidad y este sentó las bases para la adaptabilidad. La fórmula que ganó mayor relevancia y la que es de más uso general es la de Rodríguez Trujillo, que se basó en el análisis de regresión múltiple del Ministerio de Educación de Venezuela (Crawford, 1989); la variable  $L = \text{número de letras}$ ,  $W = \text{número de palabras}$ ,  $S = \text{número de oraciones}$  y  $C = \text{Comprensibilidad}$ , de esta forma obteniendo la siguiente fórmula:

$$C = 95,2 - \left(9,7 \times \frac{L}{W}\right) - \left(0,35 \times \frac{W}{S}\right)$$

Esta fórmula fue validada para el nivel de 6° hasta el uso de los *cloze* probada por los estudiantes. No es muy utilizada por falta de investigación y apoyo.

- **Perspicuidad:** Según la RAE (2014) lo define como “claro, transparente y dicho del estilo inelegible” por lo que se define a perspicuidad como la forma en la que los lectores perciben la idea del texto y cuan fácil y claro es percibir esta información. El resultado de una correcta perspicuidad en los textos es una lectura de mayor comprensión y rapidez. (Lester & Wheeler, 1955)

La fórmula de validación fue una adaptación de Rudolf Flesh de su trabajo *Facilidad de Lectura en idiomas español y francés* (Ríos, 2017). En un esfuerzo por mejorar las fórmulas matemáticas para la legibilidad en el idioma español, surge la fórmula de la perspicuidad de Francisco Szigrist Pazos el cual en su tesis doctoral cita y se basa en la de Flesh, está nueva forma analiza la comprensión lectora y se divide en 2 versiones (Rios Hernandez, 2017). Entre sus principales ventajas es que estas fórmulas pueden adaptarse a cualquier tema de diversas publicaciones además categoriza el nivel de dificultad de un texto con su

fórmula para de esta forma percibir el nivel académico del autor. Las variables para estas fórmulas son las siguientes  $P = Perspicuidad$ ,  $s = Sílabas$ ,  $p = Palabras por frase$  y  $f = Frases$ . La primera fórmula se basa en textos con palabras menores a 100 y es la siguiente:

$$P = 207 - 0,623 s - p$$

Mientras que la segunda se puede utilizar para cualquier tamaño de la muestra:

$$P = 207 - (62,3 s): p - (p: f)$$

Todas estas pautas han sido comprobadas científicamente para verificar el uso de la legibilidad en la lectura. Existen varias más basadas en distintos artículos científicos que podrían enmarcarse desde varias perspectivas diferentes, además de ser conscientes de que las TIC ya suponen un reto de accesibilidad y la función con la legibilidad supondría un gran avance para este reto.

- **Lectorabilidad:** Podemos definirla como la facilidad o dificultad que se tiene cuando se lee un texto es por esa la importancia entre la lectorabilidad y la comprensión lectora por lo que hace énfasis en los aspectos lingüísticos, léxicos y gramaticales. Las fórmulas para la lectorabilidad tienen bastante historia y todas escritas por hispano hablantes, estas fórmulas se basan en la longitud de las palabras y fue Fernández Huerta el precursor en el año 1959 y este fue quien dio paso a las fórmulas que le presidieron. Por ejemplo, el Gráfico de Fry toma en cuenta una conexión entre el número de oraciones y el número de sílabas del texto que se va a analizar, o la fórmula de Spaulding que creó una lista denominada “Density Word List” en la que proponía a una fórmula basada en la longitud de frases y la densidad de vocabulario en base a las que no se encuentran en esa lista (Ríos Hernández, 2015). Pero la fórmula general y la que se utiliza normalmente es la de Fernández, en la cual  $L = Lectorabilidad$ ,  $P = número de sílabas$  y  $F = número de frases$  cada una por cada 100 palabras.



$$L = 206,84 - 0,60 P - 1,02 F$$

Siempre se anota al principio de cada texto el número de sílabas por cada 100 palabras que se consideran. Las frases se diferencian por el punto o el punto y coma.

### 3.3. Normas de Lectura Fácil

Este proyecto se ha enfocado en el uso de normas y pautas brindadas por la CONADIS en Ecuador, pero es importante recalcar los antecedentes que dieron paso a la creación de esta y en cuales fue basada. Es importante reiterar que en 2007 fue publicada la primera adaptación a lectura fácil sobre los Derechos de las Personas con Discapacidad de las Naciones Unidas luego de haber adaptado un el libro Don Quijote de la Mancha en 2005.

#### **UNE 153101:2018 EX – Lectura Fácil: Pautas y Recomendaciones para la elaboración de documentos**

España es el país pionero en la redacción de textos a lectura fácil para la lengua española, la norma UNE 153101 EX busca garantizar el derecho el acceso a la información facilitando la accesibilidad cognitiva universal para la igualdad de oportunidades. Esta norma aplica para distintos tipos de textos escritos. Para la creación del documento son necesarias dos tipos de fases, la de adaptación y creación y cada una de ellas cuenta con su fase de validación. (Asociación Española de Normalización [UNE], 2018)

1. **Fase de adaptación:** En esta fase se adapta, diseña y maqueta la información recogiendo toda la información necesaria y relevante del documento.
2. **Fase de validación:** Esta información debe realizarse por validadores, por lo que se realiza el primer borrador del texto en Lectura Fácil y se verifica si proviene del documento original.

3. **Fase de creación:** Se realiza por el autor, diseñador y maquetador. Se recogen las indicaciones y se consensua la creación del proyecto.
4. **Fase de validación:** De igual forma debe realizarse por validadores y coordinadores, con las siguientes actividades planifica las sesiones de validación.

Luego para las pautas y recomendaciones es necesario dividirlo por secciones: orto tipografía, vocabulario y expresiones, frases y oraciones, organización del texto y el estilo, presentación del documento y complementos para textuales. Cada una de estas secciones cuentan con sus respectivas normas y pautas que se crean especialmente para tener relevancia de estas.(Asociación Española de Normalización [UNE], 2018)

## **1. Manual de Lectura Fácil y Formatos Accesibles del Ecuador**

Como se indicó anteriormente, este trabajo está basado en la propuesta del CONADIS, titulado Manual de Lectura Fácil y Formatos Accesibles y al igual que las normas UNE de España se usa principalmente para facilitar la accesibilidad cognitiva porque esta requiere mayor esfuerzo. Además, esta norma está basada en la UNE 153101:2018Ex así que de la misma forma se basa en secciones para la redacción de textos en lectura fácil: ortográfico, gramatical, léxico y de estilo.

1. **Pautas de Ortografía:** No existen adaptaciones especiales en textos en lectura fácil, las mismas reglas ortográficas que utilizando la mayúscula como regla general y varias reglas más que tienen que ver con la parte ortográfica, como la puntuación, signos ortográficos y fechas y números.
2. **Pautas de Gramática:** De la misma forma debe tenerse la corrección gramatical por lo que debe evitarse estructuras y tiempos morfosintácticos complejos para evitar la dificultad de los contenidos de texto. Se analizan las formas verbales, partes de la oración y tipos de oraciones.

3. **Pautas de Léxico:** Se analiza el vocabulario, las palabras utilizadas y los accidentes semánticos. Utilizando palabras en tiempos sencillos y siempre intentar mantener la legibilidad.
4. **Pautas de Estilo:** Ofrece la forma en que se debe presentar los textos de lectura fácil y como el apoyo visual ayudaría al soporte de está utilizando pictogramas e imágenes.

Para el desarrollo de este trabajo se utilizaron las primeras tres pautas para centrarse de manera más objetiva en la forma en que se percibe el texto y la comprensión lectora permitiendo la implementación de la lectura fácil más rápida y eficaz.

## 4. METODOLOGÍA

### 4.1. Conceptos previos y recursos utilizados

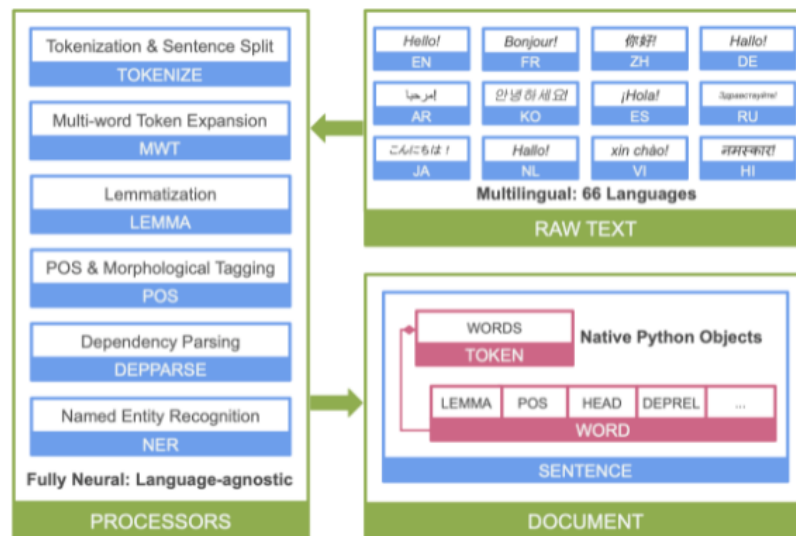
Antes de comenzar a discutir el diseño, arquitectura e implementación del módulo de adaptación de textos a un formato de lectura fácil, es necesario introducir unos conceptos fundamentales para entender el trabajo.

#### a) **Procesamiento del lenguaje natural**

El procesamiento natural de lenguaje (PLN, por sus siglas en inglés) posibilita que los dispositivos electrónicos comprendan, interpreten y manipulen el lenguaje humano con la intención de intentar simular los procesos de comunicación que tenemos las personas haciendo uso de los recursos computacionales disponibles. Esto se lleva gracias a la enorme cantidad de información que podemos encontrar hoy, la misma que está organizada en grandes conjuntos de datos estructurados, con características que nos permiten realizar algoritmos que puedan realizar tareas relacionadas al lenguaje o desarrollar modelos (Inteligencia artificial) que nos ayude a entender los procesos de comunicación que tenemos los humanos, algunas de las aplicaciones del PLN son: chatbots, extracción de información, análisis sintáctico, resúmenes de texto, etc.(Vásquez et al., 2009)

- a. *Stanza nlp*: Es un paquete de Python con distintas funciones para el análisis del lenguaje en distintos idiomas como el español. En el sitio oficial de la librería

se especifican sus funciones: “convertir una cadena que contiene texto en lenguaje humano en listas de oraciones y palabras, para generar formas base de esas palabras, sus partes del discurso y características morfológicas, para dar un análisis de dependencia de estructura sintáctica, y para reconocer entidades nombradas”. (Stanford NLP Group., 2020)



**Figura 2.** Descripción de la red neuronal NLP de Stanza.

Fuente: Stanza, <https://stanfordnlp.github.io/stanza/>

**b) Redes neuronales Linear SVM**

Las redes SVM o máquinas de soporte vectorial nos ayudan en procesos de clasificación y regresión, cuyo comportamiento se basa en la separación de los datos haciendo uso de un hiperplano cuyo comportamiento se describe igual a la ecuación de una recta, y dentro de esto hiperplano el objetivo es determinar cuál es el clasificador óptimo para la predicción de resultados, basándose en la distancia de todos los datos, a esto se le conoce como “maximal margin classifier”.(Kowalczyk, 2017)

- a. *Support vector classifier*: debido a los problemas que presenta el uso de un maximal margin classifier como la dependencia de que los datos sean perfectamente separables, se introduce el uso de un clasificador basado en hiperplano, disminuyendo así la probabilidad de un overfitting, esto se logra gracias a la permisividad de que algunos datos se encuentren en el lado incorrecto del hiperplano.(Kowalczyk, 2017)

- b. *Linear kernel*: este kernel es el más recomendado al momento de clasificar texto, este kernel es simple y se recomienda al momento de trabajar con varias características, en este caso, documentos y palabras.(Humano, n.d.)

c) **Spanish BERT (BETO)**

Dentro de los recursos y tecnologías desarrollados para el PLN nos encontramos con una biblioteca conocida como BERT, la cual brinda un modelo basado en Transformers que gracias a sus características realizan un entrenamiento bidireccional que permite mejorar los resultados.

- a. *Transformers*: Cuando se trata del uso de IA para trabajar con contenido textual y herramientas para realizar un proceso de PLN, las redes más utilizadas eran las que se conocen como redes neuronales recursivas, sin embargo, uno de los mayores problemas era la poca memoria que disponía al trabajar con textos largos, después de algunas pasadas los pesos y relaciones de los primeros términos se perdían lo que afectaba el entrenamiento. Una solución a este problema es el uso de **transformers**, la ventaja de este tipo de red neuronal es la cantidad de capas que tiene y el funcionamiento de cada una de ellas lo que mejora la capacidad de memoria para mantener los pesos de entrenamiento iniciales y así evitar los problemas que podemos encontrar cuando una sección de la oración haga referencia a algún término del inicio, el almacenamiento de pesos ayudara a que la red sepa que se está hablando de ese término en particular.(Remón, n.d.)

El principal inconveniente de esta biblioteca es que su entrenamiento esta realizado principalmente para trabajar con textos en inglés (Vásquez et al., 2009). Algunas de las soluciones que podemos encontrar para trabajar con otros idiomas como el español son variaciones que han sido desarrollados por grupos de investigación de distintas universidades e instituciones, una de ellas es un modelo desarrollado en la Universidad de Chile conocido como BETO, este modelo, cuya fuente de conocimiento viene de grandes cantidades de datos en español, como el contenido de Wikipedia, presenta funciones muy útiles a la hora de identificar relaciones entre palabras y dependencia de términos. La función que nos enfocaremos es en la predicción de términos basado en el contexto de la oración, que servirá dentro del módulo de adaptación para brindar opciones de términos que pueden resultar complejos o difíciles de entender.

d) *Expresiones Regulares*

Una expresión regular son cadenas de texto que contiene algún patrón en específico y que pueden o no contener cierto tipo de caracteres que pueden afectar una búsqueda normal de alguna palabra dentro de un texto, dentro del lenguaje Python existen bibliotecas (re) que permiten hacer uso de las expresiones regulares y que dentro del módulo de adaptación servirán para encontrar ciertos caracteres que deberá ser sustituidos o eliminados. (*Expresiones Regulares / AppExpert*, n.d.)

e) ***Web Scraping***

Esta técnica consiste básicamente en hacer una extracción de información de los contenidos que se pueden encontrar en la web, por ejemplo, sus textos, imágenes, recursos, etc. Su funcionamiento es simple, dentro del internet una búsqueda y consulta de alguna página web se realiza enviando código HTML como petición y esta petición es retornada al usuario en código HTML, entonces lo que se realiza es una limpieza de esa respuesta de manera que solo se conserva el contenido y las etiquetas y estructura HTML son ignoradas. (Profesionales et al., 2017)

## **4.2. Diseño y desarrollo del modulo**

Para el diseño del módulo de adaptación de textos a lectura fácil se procedió a trabajar con el manual de CONADIS (CONSEJO NACIONAL PARA LA IGUAL DE DISCAPACIDADES), de donde se obtiene un listado de pautas sugeridas para adaptar textos, en base al planteamiento del módulo se realiza un modelo que automatice la aplicación de estas pautas en sus distintas categorías: ortográficas, gramaticales y de léxico.

El módulo de adaptación se desarrolla haciendo uso de las librerías ya descritas, y en base a conocimientos del lenguaje en español, con esto se aplican técnicas de programación para obtener oraciones adaptadas a LF.

### **1) Pautas Ortográficas**

Dentro de este apartado se va a tratar la técnica propuesta para implementar la adaptación de pautas ortográficas dentro de un texto, la naturaleza de estas pautas nos permite reconocer 2 procesos, reemplazar contenido y eliminar contenido para realizar este proceso se realizan las siguientes acciones:

#### **a. Creación de recursos.**

Muchas de las pautas incluyen una búsqueda de contenido dentro del texto, el cual debe ser eliminado o remplazado, si bien las bibliotecas disponibles para PLN contienen ya algunas funciones para identificar ciertos términos como: conjunciones, nombres de entidades o nombres propios, algunos recursos dentro de nuestro idioma no están disponibles por ello es necesario que se disponga de recursos propios para los procesos de adaptación. Para esto se hace el uso de diccionarios para una búsqueda de clave-valor que nos permita identificar algunos términos como: refranes, tecnicismos, caracteres especiales, números, etc., los cuáles serán remplazados o eliminados según dicte la norma.

**b. Uso de Expresiones regulares**

Con los recursos disponibles para identificar los terminas a ser remplazados el uso de bibliotecas para búsqueda de expresiones regulares ayuda a identificar estos términos en el texto que va a ser adaptado.

## 2) Pautas Gramaticales

Este viene a ser considerado el tipo de pautas más importantes para adaptar nuestro texto, debido a que se centra en la forma de escritura de las oraciones y se busca llegar a la estructura más simple de nuestro idioma, una oración que contenga *sujeto + verbo + complemento* para lograr esto una solución es trabajar con redes neuronales que identifiquen ciertos tipos de oración (pasivas, coordinadas y subordinadas) y con ello se realiza un proceso de adaptación, siempre y cuando sea factible, una de las limitaciones de este módulo es el no poder trabajar con oración que tienen un sujeto tácito.

**a. Dataset**

Al no existir dentro de la comunidad un dataset que identifique el tipo de oración en lenguaje español, se procedió a crear un dataset propio, el cual está basado en contenido de material educativo como los OAs, búsquedas en internet y contenido de textos, es un dataset simple que consiste en dos atributos, oración y tipo de oración, que nos ayudara a identificar la oración y con ello se procederá a realizar el proceso de adaptación.

El dataset es un dataset muy pequeño que consiste en un total de 1000 muestras recopiladas de distintas fuentes.

**b. Preprocesamiento de datos**

Las técnicas más conocidas para realizar una limpieza de datos son, el eliminar stopwords, mantener las palabras en su forma raíz entre otras. Sin embargo, por la naturaleza de las oraciones que deben ser identificadas se realiza una adaptación a la base de datos de los stopwords, esto con el fin de evitar eliminar palabras que son importantes en las estructuras de las oraciones y que pueden ser eliminadas durante el proceso de limpieza.

*Oraciones pasivas:* Las oraciones pasivas tienen la característica de ser un sujeto que está siendo afectado por una acción y para esto mantiene verbos auxiliares como: fui, fue, fueron, estaba, esta, etc. Que dentro de las bibliotecas son considerados como stopwords, pero no deben ser eliminadas debido a que son importantes para la identificación del tipo de oración.

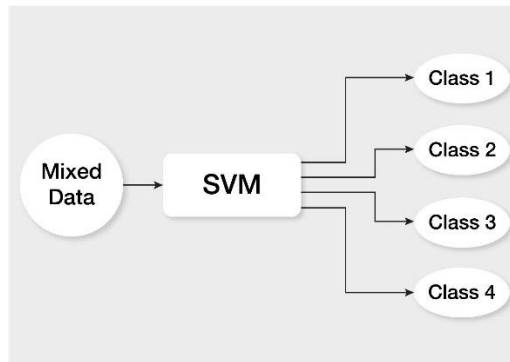
*Oraciones coordinadas:* Las oraciones coordinadas tienen la característica de reunir dos situaciones (similares o distintas) por medio de conjunciones, las cuales dentro de una bolsa de palabras son consideradas como stopwords y que deben ser eliminadas, es por ello que es importante modificar la base de conocimiento de stopwords para evitar eliminar información importante.

*Oraciones subordinadas:* Dentro de estas oraciones no existe ninguna palabra que se considere un stopwords, sin embargo, aquí es importante reconocer signos de puntuación como *comas* que nos ayuden a identificar la información extra que debe ser eliminada.

### c. Red Neuronal general

Esta red llevará a cabo un aprendizaje supervisado la arquitectura del modelo se basa en una red SVM, la cual es la más apropiada por su rapidez y sencillez para trabajar con textos.(Kowalczyk, 2017)

#### i. Diseño del modelo





**Figura 3.** Arquitectura general de una red SVM, Linear SVC. **Fuente:** Towardsdatascience, Rushikesh Pupale; <https://towardsdatascience.com/https-medium-com-pupalerushikesh-svm-f4b42800e989>

```
{'C': 1.0,  
  'class_weight': None,  
  'dual': True,  
  'fit_intercept': True,  
  'intercept_scaling': 1,  
  'loss': 'squared_hinge',  
  'max_iter': 1000,  
  'multi_class': 'ovr',  
  'penalty': 'l2',  
  'random_state': None,  
  'tol': 0.0001,  
  'verbose': 0}
```

**Figura 4.** Parámetros de entrada de una red Linear SVC. **Fuente:**Autor

## ii. Métricas de evaluación

Uno de los problemas que encontramos es la baja cantidad de datos de entrenamiento por lo tanto la precisión de predicción se ve reflejado en el siguiente valor:

```
Accuracy: 71.7949%
```

**Figura 5.** Precisión del modelo propuesto. **Fuente:** Autor

## d. Adaptar oraciones en forma pasiva

Para que una oración en forma pasiva se pueda adaptar deben existir 3 características, debe tener un verbo raíz, un agente pasivo y un agente activo, caso contrario respetando las pautas de adaptación la oración no será modificada ni adaptada.

### i. Uso de PLN

Una de las funciones que tiene la librería de Stanza PLN es que nos permite obtener las características de cada palabra dentro de la oración, su tiempo, forma, número de personas, etc. Además, permite obtener la relación que tiene con el resto de las palabras, en base a ello se realiza un análisis del verbo raíz de la oración y se obtiene cada parte de la oración, siendo comúnmente la primera parte el sujeto que recibe la acción y la segunda parte el sujeto que realiza la acción.

## **ii. Identificación de sujetos**

Como se mencionó, para realizar la adaptación se deben tener 2 sujetos dentro de la oración, una vez que se conoce los sujetos, estos son analizados para saber el número de personas que existen y en base a eso determinar la forma que el verbo va a tener. Además, se analizan los determinantes de la oración y así no pierde su estructura y su correcta forma gramatical.

## **iii. Estructura de la oración**

### **Forma original**

Determinante + agente pasivo + complementos + verbo auxiliar  
(ser o estar) + determinante + agente que realiza la acción +  
auxiliar.

### **Forma adaptada**

Determinante + agente que realiza la acción + verbo adaptado +  
agente pasivo.

## **e. Adaptar oraciones coordinadas**

Una oración coordinada es aquella donde se introducen dos situaciones (mismo contexto o no) dentro de una oración y estas son unidas por una conjunción, de manera que se busca separar esas oraciones sin restarle información. También se encuentran algunas restricciones cuando el sujeto no existe dentro de la oración.

### **i. Identificación del sujeto**

Con la ayuda de Stanza PLN, se identifica el sujeto principal de la oración, este será un recurso que servirá para las dos oraciones ya que así se evita perder información.

### **ii. Identificación de la conjunción**

En base a los recursos que se crearon al inicio se obtiene una identificación de la conjunción de la oración, esto servirá para identificar donde esta cada parte de la oración compuesta, si es necesario se introducirá el sujeto en la segunda parte.

### iii. Estructura de la oración

#### **Forma original**

Sujeto + verbo + complemento + conjunción + (Puede o no existir sujeto) + verbo + complemento

#### **Forma adaptada**

**O1.** Sujeto + verbo + complemento

**O2.** Sujeto + verbo + complemento.

### f. Adaptar oraciones subordinadas

Dentro de las oraciones subordinadas se puede identificar información extra que puede ser eliminada sin alterar la información de la oración, esto gracias a que una pauta es evitar la información extra que no brinde datos puntuales. La principal estructura a analizar es aquella que contiene información extra dentro de comas.

#### i. Identificación de información extra

Con la ayuda de las expresiones regulares podremos identificar si dentro de la oración existe un espacio con 2 comas para verificar si la información escrita dentro trata acerca del sujeto, si es el caso proceder a eliminarla

#### ii. Validación

Cuando se obtiene la información extra, se comprueba que sea información extra relacionada al sujeto, esto se logra con la librería de Stanza que nos ayuda a identificar las relaciones entre palabras.

#### iii. Estructura

#### **Forma original**

Sujeto + coma + información extra + coma + complemento

#### **Forma adaptada**

Sujeto + verbo + complemento

### 3) Pautas de Léxico

Dentro de estas pautas se puede apreciar que se centra en las palabras usadas dentro de un texto, la idea principal es evitar el uso de palabras complejas que puedan provocar dificultades para entenderlas, por ello se usa una red neuronal para identificar posibles candidatos para sustituir las palabras difíciles.

#### a. Uso de la red neuronal Beto

Como ya se mencionó, BETO es una red para PLN, que en este caso será utilizada para obtener posibles candidatos de palabras complejas basados en el contexto de una oración.

```
text = "[CLS] Para solucionar los [MASK] de Chile, el presidente debe [MASK] de inmediato. [SEP]"
Posibles opciones  $\theta$  : ['problemas', 'conflictos', 'asuntos', 'males', 'temas']
```

#### b. Obtener sinónimos, procesos de web Scraping

Para evitar que las palabras sean erróneas en base al contexto, también se hace una búsqueda de sinónimos dentro de la web (<http://www.wordreference.com/sinonimos>) una vez que se obtiene los sinónimos se hace un procedimiento de web Scraping para obtener solo los datos que nos interesan

#### c. Uso de diccionario de frecuencia de palabras

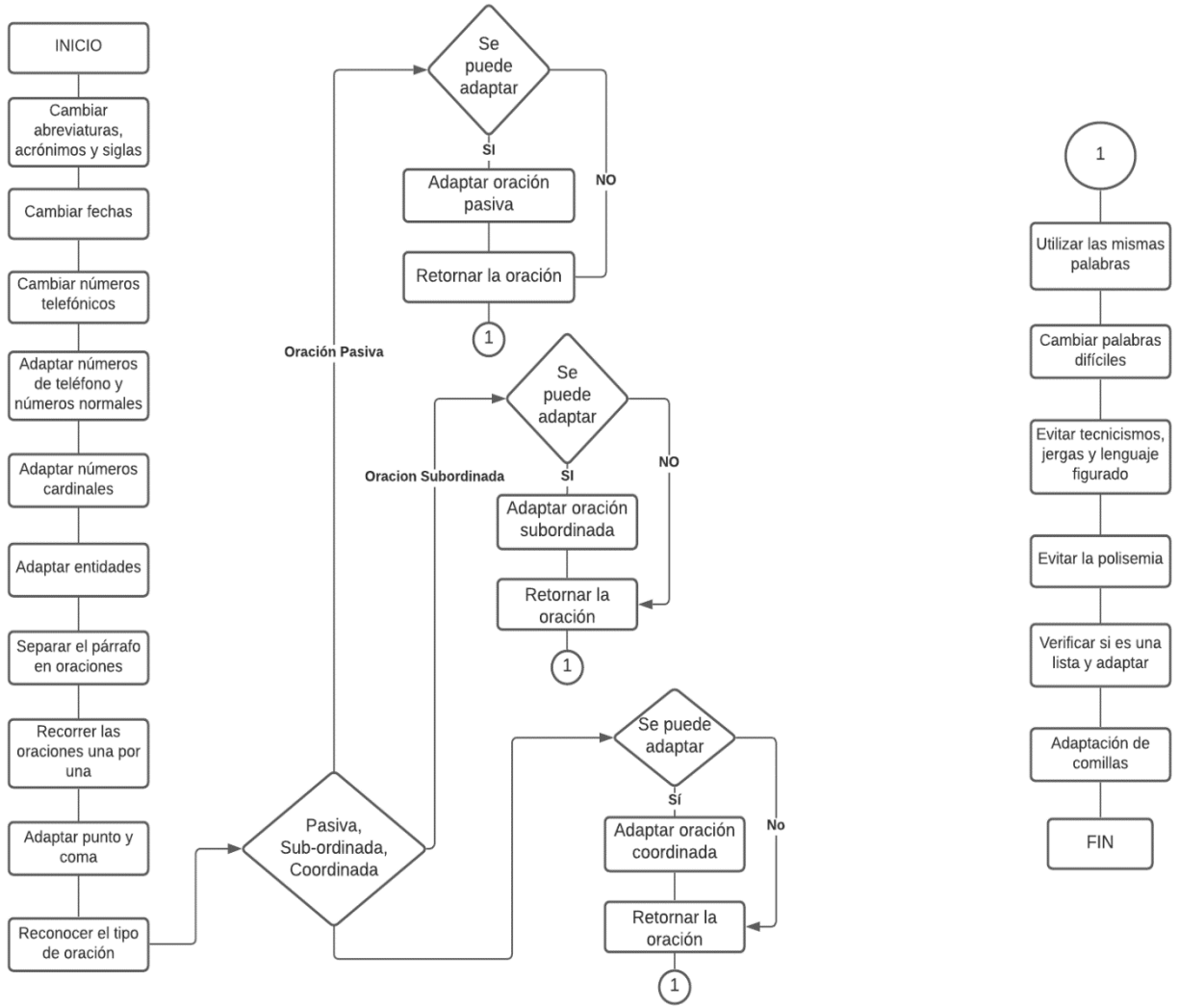
Con las opciones obtenidas a través de Beto y de la búsqueda de sinónimos se procede a identificar cuáles son las opciones que son más utilizadas en nuestro idioma, esto se logra gracias a un recurso obtenido de la RAE, en donde se tiene un listado de las palabras más utilizadas en español organizadas por términos de frecuencia. Estos datos son obtenidos de libros, textos literarios, periódicos, etc. (Isaba, 2016)

#### d. Procesos de adaptación.

La metodología que se usa es, se analiza la palabra original (su frecuencia de uso) y luego se compara con las opciones resultantes, de tal manera que, si existe una nueva palabra más usada, se remplazara.

### 4.3. Diagrama del módulo de adaptación.

**e. Diagrama estructural**

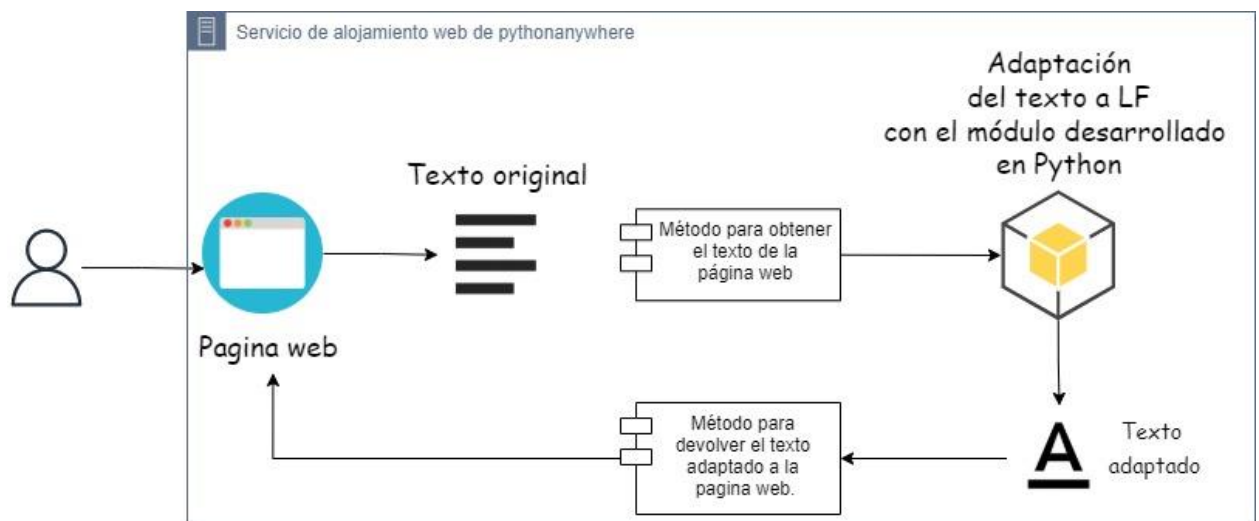


**Figura 5.** Arquitectura del módulo de adaptación. **Fuente:** Autor

#### 4.4. Despliegue en entornos definidos

##### Página web

El módulo de adaptación de texto a lectura fácil será desplegado en un servicio de alojamiento web, esto para que pueda ser probado de una manera simple dentro de una interfaz donde se puede escribir texto y obtener el mismo en un formato adaptado, el entorno escogió para realizar esto es el sitio *Pythonanywhere*, gracias a su compatibilidad con el módulo de adaptación. A continuación, la siguiente figura refleja la estructura de implementación.

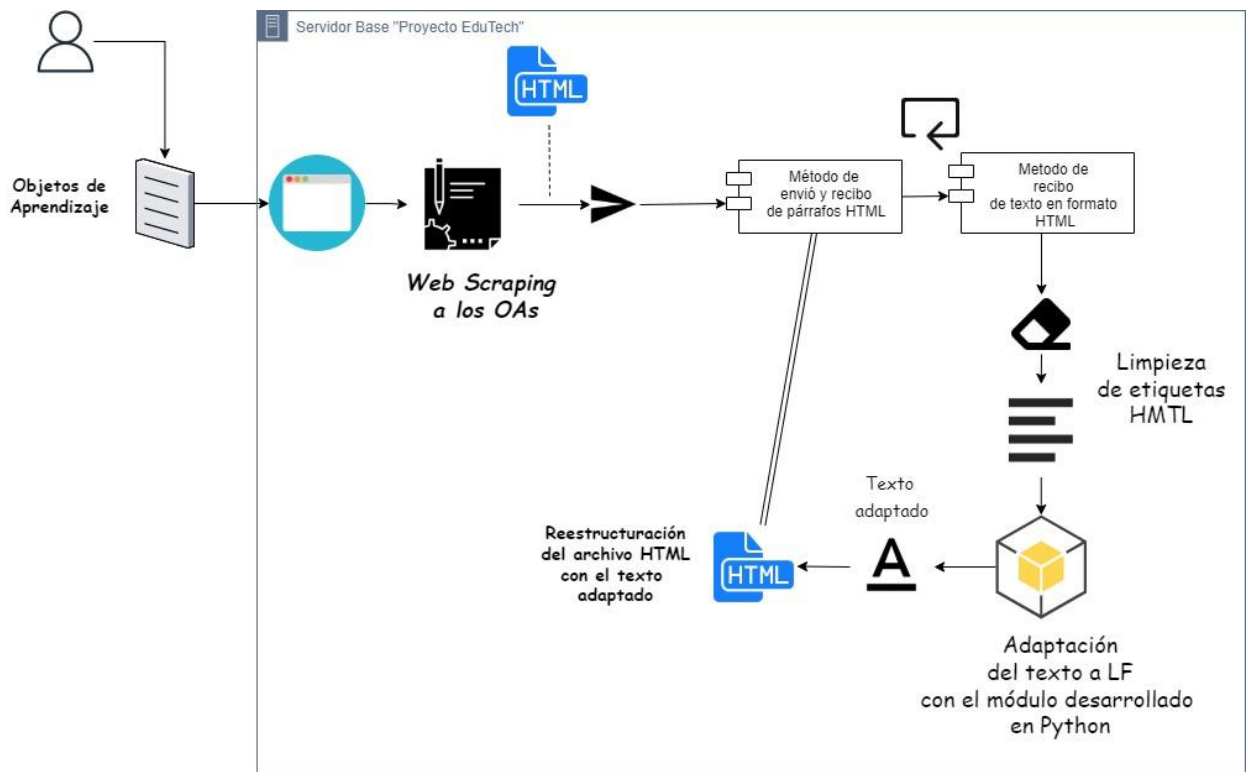


del sistema.

**Figura 6.** Arquitectura de funcionamiento de la página web con el módulo de adaptación de texto a lectura fácil. **Fuente:** Autor

##### Servidor del proyecto base “EduTech”

Uno de los entornos de producción donde el módulo será incluido es dentro del proyecto “EduTech”, dentro del cual cumple con la función de adaptar el texto extraído de los objetos de aprendizaje a un formato de Lectura fácil, la implementación de este módulo se realiza a través de métodos de conexión donde se recibe párrafo por párrafo a los cuales se les aplica una limpieza para eliminar las etiquetas HTML y de esa manera trabajar con el texto de los OAs, una vez que se obtiene el texto este es procesado por el módulo de adaptación y cuando los resultados son reestructurados en el formato HTML en el que llegó al inicio y es retornado a los otros módulos de EduTech, de esa manera es presentado en la página web del proyecto. En la siguiente figura se puede apreciar la estructura del proyecto comunicándose con el módulo de adaptación.



**Figura 7.** Arquitectura de funcionamiento y comunicación del módulo de adaptación de texto a lectura fácil con el proyecto “EduTech”. **Fuente:** Autor.

## 5. RESULTADOS Y EVALUACIONES

Los resultados de las funciones obtenidas con el módulo de adaptación se pueden apreciar en la siguiente tabla, donde se muestra las pautas que se lograron llevar a procesos de automatización.

TIPO DE PAUTA	SECCIÓN	PAUTA	COMPLETADO
<b>Ortografía</b>	Mayúsculas	Uso de la mayúscula como regla general.	SI
		Evitar palabras o frases escritas en mayúsculas.	SI
	Puntuación	Uso del punto y aparte.	SI
		Evitar el punto y seguido.	SI
		Evitar el punto y coma y los puntos suspensivos.	SI
	Signos	Uso de dos puntos especialmente para las listas.	SI
		Utilizar y respetar los signos de entonación.	SI
		Evitar corchetes y signos ortográficos.	SI
		Utilizar el guion o raya para diálogos.	SI
		Uso de comillas, explicación.	SI
	Números	Uso de números en cifra aplicando redondeo.	SI
		Redacción de números de teléfono de forma estándar.	SI



		Uso de formato de fecha.	SI
		Evitar números romanos.	SI
<b>Gramática</b>	Verbos	Evitar tiempos y modos verbales.	SI
		Uso del imperativo.	SI
		Evitar voz pasiva.	SI
		Evitar la omisión del sujeto.	NO
		Uso de aposiciones y vocativos.	SI
		Sustitución de elementos en la frase.	NO
	Oraciones	Uso de oraciones simples y cortas.	SI
		Evitar oraciones impersonales.	NO
		Evitar oraciones coordinadas y subordinadas.	SI
		Evitar formas negativas.	NO
Uso de un estilo directo.		SI	
<b>Léxico</b>	Palabras	Uso de palabras sencillas.	SI
		Evitar palabras largas o difíciles.	SI
		Repetir términos legibilidad.	NO
		Evitar abreviaturas, acrónimos y siglas.	SI
		Evitar verbos normalizados.	SI
		Uso de adverbios (-mente).	SI

		Uso de preposiciones y conjunciones.	SI
	Accidentes Semánticos	Uso de onomatopeyas.	SI
		Evitar la polisemia.	NO
		Uso de antónimos.	SI
		Evitar tecnicismos, jergas y xenismos.	SI
		Evitar lenguaje figurado.	SI
		Evitar expresiones personales.	SI
		Evitar porcentajes.	SI
		Usar el mismo sinónimo.	SI

### 5.1. Evaluaciones con OAs

Titulo	COVID19
Texto original	Texto adaptado
<p>Texto tomado de objetos de aprendizaje: “Los síntomas más habituales de la COVID-19 son la fiebre, la tos seca y el cansancio. Otros síntomas menos frecuentes que afectan a algunos pacientes son los dolores y molestias, la congestión nasal, el dolor de cabeza, la conjuntivitis, el dolor de garganta, la diarrea, la pérdida del gusto o el olfato y las erupciones cutáneas o cambios de color en los dedos de las manos o los pies. Estos síntomas suelen ser leves y comienzan gradualmente. Algunas de las personas infectadas solo presentan síntomas levisimos.</p> <p>La mayoría de las personas (alrededor del 80%) se recuperan de la enfermedad sin necesidad de tratamiento hospitalario. Alrededor de 1 de cada 5 personas que contraen la COVID19 acaba presentando un cuadro grave y experimenta dificultades para respirar. Las personas mayores y las que padecen afecciones médicas previas como hipertensión arterial, problemas cardiacos o pulmonares, diabetes o cáncer tienen más probabilidades de presentar cuadros graves. Sin embargo, cualquier persona puede contraer la COVID19 y caer gravemente enferma. Las personas de cualquier edad que tengan fiebre o tos y además respiren con dificultad, sientan dolor u opresión en el pecho o tengan dificultades para hablar o moverse deben solicitar atención médica inmediatamente. Si es posible, se recomienda llamar primero al profesional sanitario o centro médico para que estos remitan al paciente al establecimiento sanitario adecuado.”</p>	<p>Texto obtenido de la herramienta: “Los síntomas más habituales de la enfermedad son:</p> <ul style="list-style-type: none"> <li>• <u>fiebre</u></li> <li>• <u>la tos seca</u></li> <li>• <u>y el cansancio</u></li> </ul> <p>Otros síntomas menos comunes que afectan a algunos pacientes son</p> <ul style="list-style-type: none"> <li>• los dolores y molestias</li> <li>• la congestión nasal</li> <li>• el dolor de cabeza</li> <li>• la enfermedad</li> <li>• el dolor de garganta</li> <li>• la diarrea,</li> <li>• la pérdida del gusto o el olfato y las erupciones cutáneas o cambios de color en los dedos de las manos o los pies.</li> </ul> <p>Estos síntomas suelen ser menores.  Estos síntomas comienzan gradualmente.  Algunas de las personas contagiadas solo presentan síntomas muy leves.  La mayoría de las personas se recuperan de la enfermedad sin necesidad de tratamiento médico.  Alrededor de una de cada cinco personas que contraen la enfermedad acaba presentando un cuadro grave y experimenta dificultades para respirar.  Las personas mayores.  Las que padecen afecciones médicas anteriores como hipertensión arterial, problemas cardiacos o pulmonares, diabetes o cáncer tienen más probabilidades de presentar cuadros graves.  Cualquier persona puede contagiarse la enfermedad.  Caer gravemente enferma.  Las personas de cualquier edad que tengan fiebre o tos.  Además, respiren con dificultad, sientan dolor y opresión en el pecho o tengan problemas para hablar o moverse deben pedir atención médica de manera inmediata.  Si es posible, se recomienda llamar primero al profesional sanitario.  Centro médico para que estos envíen al paciente al establecimiento sanitario adecuado.”</p>

		Tema		Covid-19	
		Índice	Valor	Dificultad	
<b>Texto Original</b>	<i>Lectorabilidad</i>		57.58	Algo difícil	
	<i>Comprensibilidad</i>		37.05	Normal	
	<i>Perspicuidad</i>		52.57	Algo difícil	
	<i>Legibilidad</i>		50.82	Difícil	
<b>Texto Adaptado</b>	<i>Lectorabilidad</i>		60.53	Normal	
	<i>Comprensibilidad</i>		37.58	Normal	
	<i>Perspicuidad</i>		55.76	Normal	
	<i>Legibilidad</i>		55.66	Un poco difícil	

<b>Título</b>	<b>Álgebra de Boole</b>
<b>Texto original</b>	<b>Texto adaptado</b>

Texto obtenido de los objetos de aprendizaje: Texto obtenido de la herramienta: “George Boole (1815 - 1864) fue un famoso matemático de origen inglés que publicó un tratado sobre las leyes del pensamiento, el cual sustenta las teorías de la lógica y la probabilidad. El siguiente video elaborado por UCC Brings Boole2School (2015) al celebrar el bicentenario del nacimiento de George Boole te permitirá conocer un poco más sobre su vida.

El siguiente video hecho por UCC Brings Boole2School al celebrar el bicentenario del nacimiento de George Boole te permitirá conocer un poco más sobre su vida.

El Álgebra de Boole permitió establecer la analogía entre los símbolos algebraicos y aquellos que representan sus formas lógicas. Algunos aspectos de interés sobre ello:

El Álgebra de Boole permitió establecer la analogía entre los símbolos algebraicos y aquellos que representan sus formas lógicas. Algunos aspectos de interés sobre ello:

- El álgebra Booleana o de Boole, consiste en un método para resolver problemas de lógica usando para ello valores binarios **uno** y **cero** y a tres operadores:
  - **AND (y)**
  - **OR (ó)**
  - **y NOT (no).**
- El álgebra Booleana se emplea en la creación de computadoras, circuitos eléctricos.
- En términos generales, el álgebra de Boole se puede aplicar a cualquier sistema en que cada variable posea dos estados.
- Las computadoras modernas realizan sus operaciones.
- Almacenan la información aplicando valores binarios
- Realizan sus operaciones en base al álgebra de Boole.
- Se emplea en el diseño de circuitos electrónicos.
- Se usa ampliamente en la programación.”

- El álgebra Booleana o de Boole, consiste en un método para resolver problemas de lógica empleando para ello valores binarios 1 y 0 y a tres operadores: AND (y), OR (ó) y NOT (no).
- El álgebra Booleana se emplea en la construcción de computadoras, circuitos eléctricos, etc.
- Es fundamental destacar que los sistemas modernos trabajan a partir de lógica binaria.
- En términos generales, el álgebra de Boole se puede emplear a cualquier sistema en que cada variable posea dos estados (1 y 0).
- Las computadoras modernas realizan sus operaciones y almacenan la información empleando valores binarios (1 y 0) y realizan sus operaciones en base al álgebra de Boole.
- Se emplea en el diseño de circuitos electrónicos.
- Se emplea ampliamente en la programación (condiciones lógicas).”

	Tema		Álgebra de Boole
	Índice	Valor	Dificultad
<b>Texto Original</b>	<i>Lectorabilidad</i>	61.43	Normal
	<i>Comprensibilidad</i>	42.21	Normal
	<i>Perspiciuidad</i>	56.3	Normal
	<i>Legibilidad</i>	51.17	Un poco difícil
<b>Texto Adaptado</b>	<i>Lectorabilidad</i>	59.95	Algo difícil
	<i>Comprensibilidad</i>	41.95	Normal
	<i>Perspiciuidad</i>	55.11	Normal

	<i>Legibilidad</i>	52.4	Un poco difícil
--	--------------------	------	-----------------

## 6. CONCLUSIONES Y TRABAJO FUTURO

Durante el desarrollo del módulo, se pudo apreciar que el área del lenguaje y especialmente dentro del idioma español, tiene una gran amplitud, esto debido a las múltiples variaciones que puede tener un texto y a esto se le puede agregar que, la cantidad de datos (corpus) en idioma español que ayuden en el desarrollo de herramientas inteligentes es muy reducida en comparación de los datos que se encuentran para otros idiomas como el inglés.

Este módulo se ha enfocado en las principales estructuras gramaticales que podemos encontrar en el lenguaje español, con esto se busca decir que, cuando existan ciertos tipos de estructuras que contengan algún componente que no ha sido tratado dentro del módulo, los resultados podrían variar y no ser tan eficientes como se espera pero cuando las estructuras son las previstas, el análisis de las pruebas muestra como la herramienta es eficiente y presenta una solución eficaz para la automatización de los procesos de adaptación, sin embargo es importante recalcar que siempre se debe revisar los resultados finales y cuando sea necesario deberá existir la intervención humana, esto con el fin de corregir posibles situaciones que afecten al contenido textual.

Con esta herramienta se aprecia que se están aprovechando tecnologías que pueden resultar un gran avance en los procesos de garantizar el derecho a la accesibilidad para las personas que se sienten rezagadas y así poco a poco seguir fomentando la inclusión.

Este proyecto se puede considerar como una base para futuras versiones que mejoren sus resultados tanto en rendimiento como en aprovechamiento de recursos computacionales. Existen múltiples estructuras del lenguaje que aún deben ser consideradas para evitar inconvenientes en cualquier texto de cualquier temática, las cuales deben ser incluidas a esta herramienta. Además, al ser una herramienta netamente enfocada a la adaptación de texto y al

estar basada en las normas del CONADIS podría incluirse las normas faltantes (pautas de estilo) para llevar a cabo una mejor y más completa adaptación de los textos.

Los resultados del desarrollo indican que no todas las normas pudieron ser incluidas en el módulo por lo tanto se puede realizar una investigación más profunda e implementarlas.

## 7. REFERENCIAS

Adelbery, A. H., & Razeq, J. R. (1984). The Cloze Procedure: A Methodology for Determining the Understandability of Accounting Textbooks. *Accounting Review*, 59(1), 109–122.

Asociación Española de Normalización [UNE]. (2018). *Norma Española Experimental UNE 153101 EX. Lectura Fácil: Pautas y recomendaciones para la elaboración de documentos*. <https://www.une.org/encuentra-tu-norma/busca-tu-norma/norma?c=N0060036>

Campus, V. (2020). *El educador, un tejedor de sueños PERIÓDICO CAMPUS pp3. Ed.1.296. September 2015.*

Crawford, A. (1989). Fórmula y gráfico para determinar la comprensibilidad de textos de nivel primario en castellano. *Https://Legible.Es/Blog/Formula-De-Crawford/*.

*Expresiones regulares | AppExpert*. (n.d.). <https://docs.citrix.com/es-es/citrix-adc/current-release/appexpert/policies-and-expressions/regular-expressions.html>

Harari, I., & Amadeo, P. (2020). Accesibilidad web. *Accesibilidad Web*. <https://doi.org/10.35537/10915/105786>

Humano, G. D. E. C. (n.d.). “ *CARACTERIZACIÓN DE LA POBREZA SUPERVISADO EN EL ESTADO DE*.”

Kowalczyk, A. (2017). Support Vector Machines Succinctly, Syncfusion. *Succinctly E-Book Series*, 114. [www.syncfusion.com](http://www.syncfusion.com).

Lester, R., & Wheeler, V. D. (1955). Improving Rate of Comprehension. *The High School Journal*,



39(1), 53–58.

Ministerio de Educación. (2021). “ *Juntos Leemos .* ” 92.

Nomura, M., Nielsen, G. S., & Tronbacke, B. (2010). *Directrices para Materiales de Lectura Fácil*. 45.

Profesionales, I. D. E. A., Elva, M. I. N., & Rodríguez, C. (2017). *El Uso De Web Scraping Para*.

Remón, G. M. (n.d.). *Tfm\_Guillermo\_Marco\_Remon*.

Ríos Hernández, I. N. (2017). Un acercamiento a la legibilidad de textos relacionados con el campo de la salud. *Chasqui. Revista Latinoamericana de Comunicación*, 135, 253. <https://doi.org/10.16921/chasqui.v0i135.2892>

Ríos Hernández, I. N. (2015). *Influencias del lenguaje y origen de un lector en la comprensión de mensajes de comunicación en salud y en la formación de actitud e intención hacia la realización de una conducta preventiva*. 238.

Ríos, I. N. (2017). Un acercamiento a la legibilidad de textos relacionados con el campo de la salud. *Chasqui. Revista Latinoamericana de Comunicación*, 135, 253–273.

Tuteladas, J., Viven, Q. U. E., Centros, E. N., En, R., & Unión, L. A. (2009). *Manual de Lectura Fácil y Formatos Accesibles*. 1–4.

Vásquez, A. C., Huerta, H. V., Quispe, J. P., & Huayna, A. M. (2009). Procesamiento de lenguaje natural robusto. *Revista de Ingeniería de Sistemas e Informática*, 6(2), 45–54. <https://revistasinvestigacion.unmsm.edu.pe/index.php/sistem/article/view/5923>

Adelbery, A. H., & Razeq, J. R. (1984). The Cloze Procedure: A Methodology for Determining the Understandability of Accounting Textbooks. *Accounting Review*, 59(1), 109–122.

Asociación Española de Normalización [UNE]. (2018). *Norma Española Experimental UNE 153101 EX. Lectura Fácil: Pautas y recomendaciones para la elaboración de documentos*.

<https://www.une.org/encuentra-tu-norma/busca-tu-norma/norma?c=N0060036>

Campus, V. (2020). *El educador, un tejedor de sueños PERIÓDICO CAMPUS pp3. Ed.1.296. September 2015.*

Crawford, A. (1989). Fórmula y gráfico para determinar la comprensibilidad de textos de nivel primario en castellano. *Https://Legible.Es/Blog/Formula-De-Crawford/.*

*Expresiones regulares | AppExpert.* (n.d.). <https://docs.citrix.com/es-es/citrix-adc/current-release/appexpert/policies-and-expressions/regular-expressions.html>

Harari, I., & Amadeo, P. (2020). Accesibilidad web. *Accesibilidad Web.* <https://doi.org/10.35537/10915/105786>

Humano, G. D. E. C. (n.d.). “ *CARACTERIZACIÓN DE LA POBREZA SUPERVISADO EN EL ESTADO DE.*

Kowalczyk, A. (2017). Support Vector Machines Succinctly, Syncfusion. *Succinctly E-Book Series*, 114. [www.syncfusion.com](http://www.syncfusion.com).

Lester, R., & Wheeler, V. D. (1955). Improving Rate of Comprehension. *The High School Journal*, 39(1), 53–58.

Ministerio de Educación. (2021). “ *Juntos Leemos .*” 92.

Nomura, M., Nielsen, G. S., & Tronbacke, B. (2010). *Directrices para Materiales de Lectura Fácil.* 45.

Profesionales, I. D. E. A., Elva, M. I. N., & Rodríguez, C. (2017). *El Uso De Web Scraping Para.*

Remón, G. M. (n.d.). *Tfm\_Guillermo\_Marco\_Remon.*

Rios Hernandez, I. N. (2017). Un acercamiento a la legibilidad de textos relacionados con el campo de la salud. *Chasqui. Revista Latinoamericana de Comunicación*, 135, 253.

<https://doi.org/10.16921/chasqui.v0i135.2892>

Ríos Hernández, I. N. (2015). *Influencias del lenguaje y origen de un lector en la comprensión de mensajes de comunicación en salud y en la formación de actitud e intención hacia la realización de una conducta preventiva*. 238.

Ríos, I. N. (2017). Un acercamiento a la legibilidad de textos relacionados con el campo de la salud. *Chasqui. Revista Latinoamericana de Comunicación*, 135, 253–273.

Tuteladas, J., Viven, Q. U. E., Centros, E. N., En, R., & Unión, L. A. (2009). *Manual de Lectura Fácil y Formatos Accesibles*. 1–4.

Vásquez, A. C., Huerta, H. V., Quispe, J. P., & Huayna, A. M. (2009). Procesamiento de lenguaje natural robusto. *Revista de Ingeniería de Sistemas e Informática*, 6(2), 45–54.  
<https://revistasinvestigacion.unmsm.edu.pe/index.php/sistem/article/view/5923>