**Documento de requerimientos de software**

***Word match pro***

***Fecha: 16-04-2024***

**Tabla de contenido**

Historial de Versiones 3

Información del Proyecto 3

Aprobaciones 3

1. Propósito 4

2. Alcance del producto / Software 4

3. Referencias 4

4. Funcionalidades del producto 5

5. Clases y características de usuarios 5

6. Entorno operativo 5

7. Requerimientos funcionales 6

9.1. (Nombre de la funcionalidad 1) 6

9.2. (Nombre de la funcionalidad 2) 7

9.3. (Nombre de la funcionalidad N) 7

8. Reglas de negocio 8

9. Requerimientos de interfaces externas 9

9.1. Interfaces de usuario 9

9.2. Interfaces de hardware 9

9.3. Interfaces de software 9

9.4. Interfaces de comunicación 9

10. Requerimientos no funcionales 10

11. Otros requerimientos 11

12. Glosario 12

**Historial de Versiones**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Fecha** | **Versión** | **Autor** | **Organización** | **Descripción** |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

**Información del Proyecto**

|  |  |
| --- | --- |
| Empresa / Organización | Word match pro |
| Proyecto | Software que extrae palabras similares de dos documentos distintos |
| Fecha de preparación | 25 de marzo del 2024 |
| Cliente | Samantha ramirez sanchez |
| Patrocinador principal | Hernandez ramos michaell roy |
| Gerente / Líder de Proyecto | Mariana aleman martinez |
| Gerente / Líder de Análisis de negocio y requerimientos | Jimenez hernandez alan yunuel |

**Aprobaciones**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Nombre y Apellido** | **Cargo** | **Departamento u Organización** | **Fecha** | **Firma** |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

1. **Propósito**

Comparación de palabras clave: Identificar y resaltar las palabras o frases clave que son comunes entre los dos documentos.

Visualización de diferencias: Mostrar de manera clara las partes del texto que son diferentes entre los dos documentos, ya sea mediante resaltado, líneas de colores, o cualquier otro método visual.

Estadísticas de similitud: Proporcionar información sobre la cantidad de palabras o frases comunes, el porcentaje de similitud, y otros datos estadísticos que ayuden a comprender la relación entre los documentos.

Funciones de filtrado y búsqueda: Permitir a los usuarios filtrar o buscar palabras específicas dentro de los documentos para facilitar el análisis.

Compatibilidad con varios formatos de archivo: Ser capaz de comparar documentos en diferentes formatos, como .docx, .pdf, .txt, etc.

Facilidad de uso: Interfaz intuitiva que permita a los usuarios cargar fácilmente los documentos, ajustar configuraciones según sea necesario y comprender los resultados de la comparación.

1. **Alcance del producto / Software**

Comparación de palabras clave: Identificar y resaltar las palabras o frases clave que son comunes entre los dos documentos.

Visualización de diferencias: Mostrar de manera clara las partes del texto que son diferentes entre los dos documentos, ya sea mediante resaltado, líneas de colores, o cualquier otro método visual.

Estadísticas de similitud: Proporcionar información sobre la cantidad de palabras o frases comunes, el porcentaje de similitud, y otros datos estadísticos que ayuden a comprender la relación entre los documentos.

Funciones de filtrado y búsqueda: Permitir a los usuarios filtrar o buscar palabras específicas dentro de los documentos para facilitar el análisis.

Compatibilidad con varios formatos de archivo: Ser capaz de comparar documentos en diferentes formatos, como .docx, .pdf, .txt, etc.

Facilidad de uso: Interfaz intuitiva que permita a los usuarios cargar fácilmente los documentos, ajustar configuraciones según sea necesario y comprender los resultados de la comparación.

1. **Referencias**

1. **Funcionalidades del producto**

Comparación lado a lado: La capacidad de ver los dos documentos uno al lado del otro para una comparación visual directa.

Resaltado de diferencias: Resaltar automáticamente las palabras o frases que son diferentes entre los dos documentos para una fácil identificación.

Resumen de similitudes: Proporcionar un resumen de las similitudes entre los documentos, como el número total de palabras en común, porcentaje de similitud, etc.

Filtrado de resultados: Permitir a los usuarios filtrar las diferencias por tipo (por ejemplo, palabras añadidas, eliminadas, modificadas) para un análisis más detallado.

Navegación intuitiva: Ofrecer herramientas de navegación que permitan a los usuarios moverse rápidamente entre las diferencias identificadas.

Soporte para diferentes formatos de archivo: Ser compatible con una variedad de formatos de archivo, como .docx, .pdf, .txt, etc., para que los usuarios puedan comparar documentos en diferentes plataformas.

Exportación de resultados: Permitir a los usuarios exportar los resultados de la comparación en diferentes formatos (por ejemplo, PDF, Excel) para su posterior análisis o informes.

Personalización de configuración: Ofrecer opciones para personalizar la configuración de la comparación, como ignorar espacios en blanco, mayúsculas y minúsculas, puntuación, etc.

Compatibilidad con varios idiomas: Ser capaz de comparar documentos en diferentes idiomas y manejar adecuadamente las diferencias lingüísticas.

Seguridad y privacidad: Garantizar la seguridad y privacidad de los documentos cargados para la comparación, especialmente en el caso de documentos confidenciales o sensibles.

Estas son algunas de las funcionalidades básicas que podrías encontrar en un software de comparación de palabras de dos documentos. Dependiendo de las necesidades específicas del usuario, el software puede ofrecer funcionalidades adicionales para mejorar la experiencia de comparación y análisis.

1. **Clases y características de usuarios**

Estudiantes:

Características: Estudiantes de todos los niveles educativos que necesitan comparar sus trabajos para evitar el plagio o para revisar sus propias versiones.

Necesidades: Funcionalidades intuitivas y fáciles de usar, detección precisa de similitudes, acceso rápido y resultados claros.

Profesores y académicos:

Características: Profesores, profesionales de la educación y académicos que necesitan verificar la originalidad de los trabajos de los estudiantes o comparar diferentes versiones de documentos.

Necesidades: Herramientas avanzadas de análisis, informes detallados sobre similitudes y diferencias, capacidad para gestionar múltiples documentos y proyectos.

Profesionales y escritores:

Características: Profesionales que necesitan comparar documentos para verificar la originalidad de contenido, realizar revisiones de documentos o colaborar en proyectos de escritura.

Necesidades: Funciones de comparación precisas y detalladas, compatibilidad con diferentes formatos de archivo, herramientas de colaboración y comentarios.

Editores y revisores:

Características: Editores y revisores que necesitan comparar documentos para detectar cambios realizados durante el proceso de edición y revisión.

Necesidades: Resaltado claro de diferencias, herramientas de navegación eficientes para revisar los cambios, capacidad para aceptar o rechazar cambios individualmente.

Profesionales de la industria legal:

Características: Abogados y profesionales legales que necesitan comparar documentos legales para identificar diferencias entre versiones o analizar contratos.

Necesidades: Precisión en la detección de cambios, herramientas de filtrado para identificar cambios específicos, seguridad y cumplimiento de normativas de privacidad.

Traductores:

Características: Traductores que necesitan comparar el texto original y la traducción para garantizar la precisión y coherencia.

Necesidades: Herramientas de alineación de texto, resaltado de diferencias lingüísticas y semánticas, compatibilidad con varios idiomas.

Usuarios corporativos:

Características: Profesionales en entornos corporativos que necesitan comparar documentos para mantener la consistencia y la precisión en la comunicación escrita.

Necesidades: Integración con sistemas de gestión de documentos, seguridad de datos, capacidad para trabajar con grandes volúmenes de documentos.

1. **Entorno operativo**

Aplicaciones de escritorio:

Sistemas operativos: Windows, macOS, Linux.

Características: Estas aplicaciones se instalan en el dispositivo del usuario y funcionan de manera independiente del navegador web. Pueden ofrecer una mayor integración con el sistema operativo y un rendimiento más rápido.

Aplicaciones web:

Navegadores web: Google Chrome, Mozilla Firefox, Safari, Microsoft Edge, etc.

Características: Estas aplicaciones se ejecutan a través de un navegador web y son accesibles desde cualquier dispositivo con conexión a Internet. Son convenientes porque no requieren instalación, pero dependen de la conexión a Internet y pueden tener algunas limitaciones en cuanto a rendimiento y funcionalidades.

Aplicaciones móviles:

Plataformas móviles: iOS, Android.

Características: Estas aplicaciones están diseñadas para dispositivos móviles como teléfonos inteligentes y tabletas. Ofrecen una experiencia optimizada para pantallas táctiles y pueden integrarse con otras aplicaciones y servicios del dispositivo.

Aplicaciones integradas en suites de productividad:

Ejemplos: Microsoft Word, Google Docs, LibreOffice.

Características: Algunas suites de productividad incluyen herramientas integradas para comparar documentos. Estas herramientas suelen ser parte de un paquete más amplio de aplicaciones de productividad y pueden ofrecer una integración más fluida con el flujo de trabajo de creación de documentos.

1. **Requerimientos funcionales**

Carga de documentos: Permitir a los usuarios cargar dos documentos en diferentes formatos, como .docx, .pdf, .txt, etc.

Extracción de texto: Extraer el contenido textual de los documentos cargados para su comparación.

Comparación de palabras: Identificar y comparar las palabras presentes en ambos documentos.

Resaltado de similitudes: Resaltar las palabras que son idénticas o similares en ambos documentos.

Detección de diferencias: Identificar las palabras que son diferentes entre los dos documentos y resaltar estas diferencias de manera clara.

Filtrado de resultados: Permitir a los usuarios filtrar las palabras por tipo de diferencia (añadidas, eliminadas, modificadas) para un análisis más detallado.

Visualización de resultados: Mostrar los resultados de la comparación de manera clara y comprensible, preferiblemente en una interfaz fácil de entender.

Exportación de resultados: Ofrecer la opción de exportar los resultados de la comparación en diferentes formatos (por ejemplo, PDF, Excel) para su posterior análisis o informes.

Navegación entre diferencias: Permitir a los usuarios navegar fácilmente entre las diferencias identificadas para revisarlas de manera individual.

Personalización de configuración: Permitir a los usuarios ajustar configuraciones como la sensibilidad de la comparación, la exclusión de ciertas palabras o caracteres, etc.

Compatibilidad con varios idiomas: Ser capaz de comparar documentos en diferentes idiomas y manejar adecuadamente las diferencias lingüísticas.

Seguridad y privacidad: Garantizar la seguridad y privacidad de los documentos cargados para la comparación, especialmente en el caso de documentos confidenciales o sensibles.

**7.1 (Nombre de la funcionalidad 1)**

Funcionalidad: Carga de documentos

Descripción:

Esta funcionalidad permite a los usuarios cargar dos documentos para su comparación. Los usuarios pueden seleccionar los documentos que desean comparar desde su dispositivo local o desde una ubicación en línea. Una vez que los documentos están cargados, el software puede extraer el texto de cada documento y prepararlo para la comparación.

Acciones y características asociadas:

Interfaz de usuario para seleccionar los documentos desde el dispositivo local o ubicación en línea.

Opción para cargar documentos en diferentes formatos como .docx, .pdf, .txt, entre otros.

Indicadores visuales para mostrar el progreso de carga de documentos.

Verificación de integridad de los documentos para garantizar que se carguen correctamente.

Mensajes de confirmación una vez que los documentos se hayan cargado con éxito.

Capacidad para cargar múltiples pares de documentos simultáneamente, si es necesario.

Beneficios:

La carga de documentos es la primera acción necesaria para realizar una comparación entre ellos. Esta funcionalidad proporciona a los usuarios la capacidad de iniciar el proceso de comparación de manera rápida y sencilla, lo que les permite analizar y entender las diferencias y similitudes entre los documentos seleccionados.

**7.2 (Nombre de la funcionalidad 2)**

Funcionalidad: Extracción de texto

Descripción:

Después de cargar los documentos, esta funcionalidad permite al software extraer el texto contenido en cada documento de manera precisa. El proceso de extracción de texto es fundamental para poder comparar de manera efectiva las palabras, frases y estructuras de ambos documentos.

Acciones y características asociadas:

Proceso automatizado para extraer el texto de los documentos cargados.

Reconocimiento de diferentes tipos de formatos de archivo para garantizar una extracción precisa del texto.

Conversión de imágenes o contenido no textual en los documentos en texto legible, si es aplicable.

Verificación de la integridad del texto extraído para asegurar que se capture todo el contenido correctamente.

Manejo de diferentes idiomas y caracteres especiales durante el proceso de extracción de texto.

Capacidad para manejar documentos largos o con formatos complejos de manera eficiente.

Beneficios:

La extracción de texto es esencial para poder realizar una comparación detallada y precisa entre los documentos cargados. Al convertir el contenido en texto legible, el software puede analizar y identificar las similitudes y diferencias entre los documentos de una manera más efectiva, lo que proporciona a los usuarios información valiosa para su análisis y toma de decisiones.

**7.3 (Nombre de la funcionalidad N)**

Funcionalidad: Identificación de similitudes y diferencias

Descripción:

Esta funcionalidad permite al software comparar el contenido de los dos documentos y identificar tanto las similitudes como las diferencias entre ellos. Es fundamental para entender cómo se relacionan los documentos y qué partes son iguales o distintas.

Acciones y características asociadas:

Comparación palabra por palabra, frase por frase o párrafo por párrafo entre los documentos.

Algoritmos de comparación que determinan la similitud entre las partes del texto.

Resaltado o marcado visual de las similitudes y diferencias encontradas.

Identificación de palabras, frases o párrafos exactamente iguales.

Detección de similitudes parciales o aproximadas, como palabras con ortografía similar o sinónimos.

Generación de un informe o resumen que enumere las similitudes y diferencias encontradas.

Beneficios:

Esta funcionalidad brinda a los usuarios una visión clara de cómo se relacionan los documentos entre sí. Les permite identificar fácilmente qué partes son idénticas, qué contenido ha sido modificado o añadido, y dónde se encuentran las discrepancias entre los documentos. Esto es fundamental para una revisión precisa y una comprensión completa de la relación entre los documentos comparados.

1. **Reglas de negocio**

Preprocesamiento del texto: Antes de comparar las palabras, se deben realizar operaciones de preprocesamiento en el texto para normalizarlo, como eliminar signos de puntuación, convertir todas las palabras a minúsculas y manejar palabras compuestas o formas verbales.

Tokenización: Dividir el texto en palabras individuales o tokens para facilitar la comparación palabra por palabra.

Extracción de palabras clave: Identificar palabras clave o términos significativos dentro del texto que puedan ser relevantes para la comparación.

Exclusión de palabras comunes: Excluir palabras comunes como artículos, preposiciones y conjunciones que no contribuyen significativamente a la comparación de los documentos.

Consideración de sinónimos: Considerar sinónimos o términos relacionados al comparar palabras, de modo que palabras similares pero no idénticas se reconozcan como tales.

Levenshtein distance: Utilizar la distancia de Levenshtein u otros algoritmos de edición de distancia para calcular la similitud entre palabras, teniendo en cuenta la cantidad mínima de ediciones (inserciones, eliminaciones o sustituciones) necesarias para transformar una palabra en otra.

Umbral de similitud: Establecer un umbral de similitud para determinar cuándo dos palabras se consideran lo suficientemente similares como para ser identificadas como tales.

Contexto de palabras: Considerar el contexto en el que aparecen las palabras dentro del texto para determinar su significado y relevancia en la comparación.

Manejo de errores ortográficos: Implementar mecanismos para corregir o manejar errores ortográficos en las palabras durante el proceso de comparación.

Evaluación de la frecuencia: Considerar la frecuencia con la que aparecen las palabras similares en los documentos para determinar su importancia relativa en la comparación.

Escalabilidad: Diseñar el software para que pueda manejar eficientemente grandes volúmenes de texto y documentos para una comparación rápida y precisa.

1. **Requerimientos de interfaces externas**
2. **Interfaces de usuario**

Interfaz de carga de documentos:

Permite a los usuarios cargar los dos documentos que desean comparar desde su dispositivo local o desde una ubicación en línea.

Proporciona indicadores visuales del progreso de carga de los documentos.

Permite cargar múltiples pares de documentos simultáneamente, si es necesario.

Interfaz de visualización de documentos:

Muestra una vista previa del contenido de los documentos cargados para que los usuarios puedan verificar que los documentos correctos se han seleccionado.

Permite a los usuarios navegar por el contenido de los documentos para revisarlo antes de la comparación.

Interfaz de configuración de comparación:

Permite a los usuarios ajustar las configuraciones de comparación, como el umbral de similitud o la exclusión de palabras comunes.

Proporciona opciones para seleccionar el tipo de similitud que se desea identificar, como palabras exactas, similitudes parciales o sinónimos.

Interfaz de resultados de comparación:

Muestra los resultados de la comparación de manera clara y comprensible.

Resalta las palabras similares o idénticas encontradas en ambos documentos.

Proporciona una lista detallada de las similitudes y diferencias identificadas, junto con el contexto en el que aparecen las palabras.

Interfaz de navegación entre similitudes:

Permite a los usuarios navegar fácilmente entre las palabras similares identificadas para revisarlas de manera individual.

Proporciona herramientas de navegación para desplazarse por el documento y encontrar todas las instancias de palabras similares.

Interfaz de exportación de resultados:

Ofrece la opción de exportar los resultados de la comparación en diferentes formatos, como PDF o Excel, para su posterior análisis o informes.

Permite a los usuarios guardar los resultados de la comparación o compartirlos con otros usuarios según sea necesario.

Interfaz de ayuda y soporte:

Proporciona acceso a documentación detallada y recursos de ayuda para guiar a los usuarios en el uso del software.

Ofrece opciones de contacto con el soporte técnico en caso de problemas o preguntas.

**Interfaces de hardware**

Dispositivos de entrada estándar: Los usuarios interactúan con el software principalmente a través de dispositivos de entrada estándar, como teclados y ratones para la interacción con la interfaz de usuario.

Pantalla y monitor: La visualización de la interfaz de usuario y los resultados de la comparación se realiza en un monitor o pantalla estándar de la computadora.

Sistema informático: El software se ejecuta en un sistema informático estándar, que puede ser una computadora de escritorio, una laptop, una tableta o incluso un teléfono inteligente, dependiendo de la plataforma en la que esté disponible el software.

Conexión a Internet (si es aplicable): Si el software tiene características que requieren conexión a Internet, como la carga de documentos desde almacenamiento en la nube o la obtención de actualizaciones, se requiere una conexión a Internet adecuada.

Requisitos de sistema mínimos: El hardware debe cumplir con los requisitos de sistema mínimos especificados por el software para un rendimiento óptimo. Esto puede incluir procesador, memoria RAM, espacio de almacenamiento y otros requisitos de hardware.

Periféricos adicionales (opcional): Dependiendo de las necesidades específicas del usuario y del entorno de uso, pueden requerirse periféricos adicionales, como escáneres para digitalizar documentos en papel para su comparación, o dispositivos de almacenamiento externo para cargar documentos desde medios extraíbles.

1. **Interfaces de software**

Panel de carga de documentos:

En la parte superior o en un área designada de la interfaz, habría un panel donde los usuarios pueden cargar los dos documentos que desean comparar. Puede haber botones de carga o una función de arrastrar y soltar para simplificar el proceso.

Configuración de la comparación:

A continuación del panel de carga de documentos, los usuarios pueden encontrar opciones para configurar los parámetros de la comparación. Esto puede incluir ajustes como el umbral de similitud, la exclusión de palabras comunes, la sensibilidad a las diferencias ortográficas, entre otros.

Visualización de documentos:

Una vez cargados los documentos y configurados los parámetros, el contenido de ambos documentos se muestra en la interfaz. Puede haber paneles separados para cada documento o una vista dividida que muestre ambos documentos simultáneamente. Esta visualización permite a los usuarios revisar el contenido y asegurarse de que los documentos correctos estén siendo comparados.

Resultados de la comparación:

Después de realizar la comparación, los resultados se muestran en un área dedicada de la interfaz. Aquí se resaltan las palabras similares o idénticas encontradas en ambos documentos. Los usuarios pueden navegar por los resultados y explorar las similitudes y diferencias identificadas.

Herramientas de navegación y manipulación:

La interfaz puede incluir herramientas para que los usuarios naveguen por los resultados de la comparación, como botones de siguiente y anterior para desplazarse entre las palabras similares identificadas. También puede haber opciones para realizar acciones específicas, como resaltar o desmarcar palabras seleccionadas.

Exportación y guardado de resultados:

Para facilitar el análisis posterior o el intercambio de resultados, el programa puede ofrecer opciones para exportar los resultados de la comparación en diferentes formatos, como PDF, CSV o texto plano. También puede permitir a los usuarios guardar los resultados en el formato del programa para su referencia futura.

Ayuda y soporte:

La interfaz puede incluir enlaces a recursos de ayuda, documentación y soporte técnico para que los usuarios puedan obtener más información sobre el funcionamiento del programa o resolver cualquier problema que puedan encontrar.

1. **Interfaces de comunicación**

Interfaz de usuario (UI):

La interfaz de usuario proporciona el medio principal de comunicación entre el usuario y el software. Esto incluye elementos visuales como botones, menús desplegables, campos de entrada de texto y paneles de resultados. La UI permite al usuario interactuar con el software cargando documentos, configurando opciones de comparación y revisando los resultados.

Interacción con el usuario:

La interfaz de usuario debe ser intuitiva y fácil de usar, con indicaciones claras sobre cómo utilizar las diferentes funciones del software. Esto puede incluir mensajes informativos, sugerencias de herramientas y guías paso a paso para ayudar al usuario a navegar por el proceso de comparación de documentos.

Comunicación entre componentes internos:

El software puede estar compuesto por varios componentes internos que colaboran entre sí para realizar la extracción y comparación de palabras. La interfaz de comunicación entre estos componentes asegura que la información se transmita de manera eficiente y precisa. Esto puede implicar el uso de interfaces de programación de aplicaciones (API), eventos de sistema, colas de mensajes u otros mecanismos de comunicación interna.

Gestión de errores y notificaciones:

La interfaz de comunicación debe manejar adecuadamente los errores y las notificaciones durante el proceso de comparación de documentos. Esto puede incluir la visualización de mensajes de error en la interfaz de usuario cuando ocurren problemas, así como la generación de alertas o registros de eventos para los administradores del sistema.

Feedback del sistema:

El software puede proporcionar feedback al usuario durante el proceso de comparación para informar sobre el progreso y el estado de la operación. Esto puede incluir barras de progreso, iconos de estado y mensajes de confirmación para indicar que una tarea se ha completado con éxito.

Compatibilidad con múltiples plataformas:

La interfaz de comunicación debe ser compatible con diferentes sistemas operativos y dispositivos para garantizar que el software pueda ser utilizado en una variedad de entornos. Esto puede implicar el diseño de interfaces de usuario responsivas que se adapten a diferentes tamaños de pantalla y dispositivos de entrada.

1. **Requerimientos no funcionales**

Eficiencia: El software debe ser eficiente en términos de tiempo y recursos computacionales para procesar grandes volúmenes de texto y realizar la comparación de manera rápida y sin retrasos significativos.

Precisión: Se espera que el software identifique correctamente las palabras similares entre los documentos y minimice los falsos positivos y negativos para proporcionar resultados precisos y confiables.

Escalabilidad: El software debe ser capaz de manejar un alto volumen de documentos y usuarios simultáneos, escalando adecuadamente su rendimiento según sea necesario.

Robustez: El software debe ser robusto y resistente a fallos, siendo capaz de manejar situaciones inesperadas como errores de entrada de usuario, documentos mal formateados o interrupciones de red sin provocar la pérdida de datos o la interrupción del servicio.

Seguridad: El software debe garantizar la seguridad y la privacidad de los documentos cargados y los resultados de la comparación, utilizando medidas adecuadas de cifrado, autenticación y control de acceso.

Usabilidad: La interfaz de usuario del software debe ser intuitiva y fácil de usar, con un diseño claro y una navegación sencilla que permita a los usuarios cargar documentos, configurar la comparación y revisar los resultados de manera eficiente.

Adaptabilidad: El software debe ser adaptable a diferentes entornos de usuario, incluyendo soporte para varios idiomas, formatos de archivo y plataformas de sistema operativo.

Mantenibilidad: El software debe ser fácil de mantener y actualizar, con un código limpio y bien documentado que facilite la identificación y corrección de errores, así como la introducción de nuevas funcionalidades en el futuro.

Interoperabilidad: El software debe ser compatible con otros sistemas y aplicaciones utilizados en el entorno del usuario, permitiendo la integración con herramientas de terceros cuando sea necesario.

Cumplimiento normativo: El software debe cumplir con las regulaciones y normativas aplicables en cuanto a la protección de datos, la privacidad del usuario y otros aspectos relevantes para su uso en un determinado contexto o industria.

1. **Glosario**

Extracción de Texto: Proceso mediante el cual el software obtiene el contenido textual de un documento en un formato legible y procesable.

Comparación de Documentos: Proceso de analizar dos documentos para identificar similitudes y diferencias entre su contenido.

Palabras Similares: Términos o palabras que comparten similitudes en su ortografía, significado o contexto.

Similitud de Palabras: Grado de parecido entre dos palabras, que puede basarse en la ortografía, la pronunciación, el significado o el contexto.

Algoritmo de Comparación: Conjunto de reglas y procedimientos utilizados por el software para analizar y comparar el contenido de los documentos.

Umbral de Similitud: Valor establecido por el usuario que determina cuándo dos palabras se consideran lo suficientemente similares como para ser identificadas como tales.

Resaltado de Similitudes: Técnica que utiliza el software para destacar las palabras similares encontradas en los documentos para una fácil identificación por parte del usuario.

Exclusión de Palabras Comunes: Proceso mediante el cual el software elimina palabras comunes como artículos, preposiciones y conjunciones durante la comparación para enfocarse en términos significativos.

Procesamiento de Lenguaje Natural (PLN): Campo de la inteligencia artificial que se centra en la interacción entre las computadoras y el lenguaje humano, utilizado en la extracción y comparación de palabras en los documentos.

Frecuencia de Palabras: Número de veces que una palabra específica aparece en un documento, utilizado para determinar su relevancia en la comparación.

Algoritmo de Levenshtein: Método utilizado para calcular la distancia entre dos cadenas de texto, que puede ser útil para identificar palabras similares mediante la medición de la cantidad de operaciones necesarias para transformar una palabra en otra.

Tokenización: Proceso de dividir el texto en unidades más pequeñas, como palabras o frases, para facilitar su análisis y comparación.

Contexto de Palabras: Información adicional alrededor de una palabra específica, como las frases o párrafos en los que aparece, que puede ser útil para comprender su uso y significado.

Interfaz de Usuario (UI): Entorno visual y de interacción a través del cual el usuario interactúa con el software para cargar documentos, configurar opciones y revisar los resultados de la comparación.

Requerimientos Funcionales y No Funcionales: Especificaciones detalladas de las funciones y características que debe cumplir el software, así como los aspectos de rendimiento, seguridad y usabilidad que deben ser considerados.