**Documento de requerimientos de software**

***ConcidenciasPDF***

***Fecha: 12/04/2024***

**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** |
| 09/04/2024 | 1 | Luis López | Clase de ingeniería de software | Se agrega software para extracción de textos de un PDF |
| 12/04/2024 | 2 | Luis López | Clase de ingeniería de software | Se agrega la función de crear crear diccionario de datos |

**Información del Proyecto**

|  |  |
| --- | --- |
| Empresa / Organización | Clase de ingeniería de software |
| Proyecto | CoincidenciasPDF |
| Fecha de preparación | 06 al 12 de abril del 2024 |
| Cliente | Unitec Atizapan |
| Patrocinador principal | Unitec Atizapan |
| Gerente / Líder de Proyecto | Luis López |
| Gerente / Líder de Análisis de negocio y requerimientos | Luis López |

**Aprobaciones**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Nombre y Apellido** | **Cargo** | **Departamento u Organización** | **Fecha** | **Firma** |
| Gerardo De la Rosa Rivera | CEO | Ingeniería en sistemas computacionales de Unitec | 12/04/2024 |  |
| William Shakespeare | Gerente de diseño de proyectos | Ingeniería en sistemas computacionales de Unitec | 12/04/2024 |  |
| Víctor Hugo | Gerente de finanzas | Finanzas | 12/04/2024 |  |
| Oscar Wilde | Gerente de calidad | Ingeniería en sistemas computacionales de Unitec | 12/04/2024 |  |

1. **Propósito**

Crear un software que sea capaz de encontrar coincidencias en el texto de un PDF a la hora de compararlo con la lista de palabras o frases que provea el usuario.

1. **Alcance del producto / Software**

* Extrae el texto de un PDF.
* Permite que el usuario suministre una lista de palabras a frases.
* Encuentra las coincidencias entre el PDF y la lista de palabras.

1. **Referencias**

UML [Welcome To UML Web Site!](https://www.uml.org/)

using System [System Namespace | Microsoft Learn](https://learn.microsoft.com/en-us/dotnet/api/system?view=net-8.0)

using System.Collections.Generic [System.Collections.Generic Namespace | Microsoft Learn](https://learn.microsoft.com/en-us/dotnet/api/system.collections.generic?view=net-8.0)

using System.Link [System.Linq Namespace | Microsoft Learn](https://learn.microsoft.com/en-us/dotnet/api/system.linq?view=net-8.0)

using System.Text [System.Text Namespace | Microsoft Learn](https://learn.microsoft.com/en-us/dotnet/api/system.text?view=net-8.0)

using Threading.Task [System.Threading.Tasks Namespace | Microsoft Learn](https://learn.microsoft.com/en-us/dotnet/api/system.threading.tasks?view=net-8.0)

using iTextSharp.text.pdf; [iTextSharp | iText PDF](https://itextpdf.com/products/itextsharp)

Tutorial para generar el extractor de PDF [How to extract text from PDF file in C# (youtube.com)](https://www.youtube.com/watch?v=y6s2mLpYfMc)

Tutorial para generar el diccionario [Trabajando con diccionarios en c#, ejemplos | Colecciones y Arreglos | Programando en C# (youtube.com)](https://www.youtube.com/watch?v=YGTGYysGlUQ)

1. **Funcionalidades del producto**

**Extraer PDF**

**Requerimientos**

* **Tener un PDF**
* **Que el PDF contenga texto**

**Diccionario Palabras**

**Requerimientos**

* **Que no sean palabras compuestas separadas por espacios (Hot dog, Llave inglesa, Corto circuito)**
* **Que las palabras estén escritas correctamente**

**Diccionario de Frases y palabras compuestas**

**Requerimientos**

* **Que las palabras o frases sean lo más cortas posible**

**Comparar coincidencias**

**Requerimientos**

* **Tener cargado un PDF**
* **Tener cargado un diccionario con palabra o frases**

1. **Clases y características de usuarios**

**Usuario que solo quiere extraer el texto del PDF:**

**Para este usuario solo la primer función ExtraerPDF es relevante.**

**Usuario que quiere encontrar coincidencias:**

**Para este usuario todas las funciones del software son relevantes.**

1. **Entorno operativo**

**Este software esta pensado para funcionar en Windows 7 y versiones posteriores. Requiere de un equipo capaz de trabajar con C# con la biblioteca iText para C#.**

1. **Requerimientos funcionales**

**7.1 Extraer PDF**

**Requerimientos**

* **Tener un PDF**
* **Que el PDF contenga texto**

**Prioridad alta**

**Esta funcionalidad trata acerca de extraer el texto de un PDF suministrado por el usuario.**

**7.2 Diccionario Palabras**

**Requerimientos**

* **Que no sean palabras compuestas separadas por espacios (Hot dog, Llave inglesa, Corto circuito)**
* **Que las palabras estén escritas correctamente**

**Prioridad Alta**

**Es que el usuario cree un diccionario con todas las palabras que quiere encontrar en el PDF.**

**Diccionario de Frases y palabras compuestas**

**Requerimientos**

* **Que las palabras o frases sean lo más cortas posible**

**Prioridad alta**

**Es que el usuario cree un diccionario con todas las frases que quiere encontrar en el PDF.**

**Comparar coincidencias**

**Requerimientos**

* **Tener cargado un PDF**
* **Tener cargado un diccionario con palabra o frases**

**Prioridad Alta**

**Es que el usuario pueda encontrar las coincidencias entre el diccionario y el PDF que suministro**

1. **Reglas de negocio**

**No se pueden extraer texto contenidos en imágenes.**

**No puede entender lenguajes no conocidos, inventados por el usuario o encriptaciones.**

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

* **No se exigieron requerimientos gráficos, o visuales. Sin embargo, el Software debe ser entendible.**

1. **Interfaces de hardware**

* **Se requiere un ordenador con Windows 7 o superior**

1. **Interfaces de software**

**Las bibliotecas incluidas en C # no son suficientes de tiene que agregar bibliotecas de iText en la versión compatible con este lenguaje. Se puede ejecutar el software en Windows exclusivamente para que la interacción con el software funcione correctamente.**

1. **Interfaces de comunicación**

**No requiere interfaces de comunicación para funcionar.**

1. **Requerimientos no funcionales**

**Que el Software sea entendible para personas con capacidades básicas de programación**

1. **Glosario**

C#: Lenguaje de programación de alto nivel orientado a objetos. Usado para crear programas de ordenador, aplicaciones móviles y aplicaciones en la nube,

Bibliotecas: Software prefabricado y establecido con reglas, listo para ser usado por otros programadores.

Funcionalidad: Cumplir requisitos y satisfacer necesidades en condiciones específicas.