

PERFIL DEL PROYECTO INTEGRADOR.

Cristhian Augusto Romero Mesías

Ingeniería de Software I

I. TEMA.

Prototipo transcriptor universal multiformato.

II. PROBLEMA.

PLANTEAMIENTO DEL PROBLEMA.

Actualmente la tecnología está cada vez abarcando nuestro estilo de vida, sobre todo en la forma de educarnos y sustentarnos. Cada vez existen más herramientas para facilitarnos el trabajo de hacer cosas manuales, tales como la transcripción, además hay situaciones en que las personas no tienen las mismas capacidades que todos, como una discapacidad. La discapacidad auditiva afecta al 5% de la población, y habrá momentos que no podrán contar con subtítulos en cada audio o video, o un estudiante no podrá transcribir un libro de cantidades grandes de páginas por que no vale la pena.

Por eso se va a desarrollar un software que ayude a todas esas personas, permitirá sacar lo que hablen en un audio o video, no importa el formato, además de poder transcribir lo mejor posible alguna foto o libro que sea imposible extraerlo.

Para analizar realmente el impacto que tendrá el software hay que entrevistar a distintos grupos en común con este problema, como estudiantes y personas que tendrían un beneficio debido a alguna limitante. Para tener resultados exactos y precisos es necesario usar el **método cuantitativo**.

FORMULACIÓN DEL PROBLEMA.

¿Cómo facilitar la transcripción de cualquier archivo multimedia para propósitos educativos y de ayuda social?

DELIMITACIÓN DEL PROBLEMA.

Objeto de estudio. – Computadora con compilador de Python.

Campo de acción. – OCR.

Lugar. – Ecuador - Quito

Tiempo. - 6 meses

III. OBJETIVOS

OBJETIVO GENERAL.

- Implementar un prototipo de un transcriptor de archivos multimedia que acepte cualquier formato y sea sencillo y práctico de utilizar.

OBJETIVOS ESPECÍFICOS.

- Diseñar la propuesta sobre el proyecto, tomando en cuenta que es para la población objetivo.
- Mostrar y hacer válida la propuesta planteada para el proyecto.
- Diseñar y confirmar la funcionalidad del software.

IV. OBJETO DE INVESTIGACION Y CAMPO DE ACCION.

OBJETO DE ESTUDIO.

En este caso, vamos a diseñar el prototipo mediante el lenguaje de programación Python, el cual el compilador requiere de forma recomendada un computador con un sistema operativo que lo soporte, para la cual usaremos Windows 10.

Ya que instalar Python es gratuito, es necesario tomar en cuenta que para simplificar gran parte del desarrollo del código se necesitarán instalar librerías, por ejemplo una de ellas es poppler, una librería que convierte en imágenes cada página de un archivo PDF. Por lo tanto también vamos a necesitar de PIP, el encargado de instalar las librerías y asignarlas para Python, la cual ya viene incluida en el instalador oficial.

CAMPO DE ACCION.

El punto fuerte de nuestro software será el OCR, del inglés Optical Character Recognition, es un proceso dirigido a la digitalización de textos, los cuales identifican automáticamente a partir de una imagen símbolos o caracteres que pertenecen a un determinado alfabeto, para luego almacenarlos en forma de datos. Así podremos interactuar con estos mediante un programa de edición de texto o similar.

Al mismo tiempo se usará un Voice Recognition, el cual se va a encargar de leer un archivo que contenga audio, como un podcast o un video, y va a extraer todas las palabras que salgan para transformarlas a texto que pueda ser editable.

V. IDEA A DEFENDER

Este software de transcripción de texto será una herramienta que ayudará a una gran cantidad de personas con la necesidad de lograr tener más conocimiento y saber lo que no pueden oír.

VI. METODOLOGIA A EMPLEAR.

Como metodología para el desarrollo de un software, se va a utilizar el de **SCRUM**. Debido a que, para lograr el desarrollo de los objetivos, tendremos que avanzar por pasos y en orden, donde cada etapa se evaluará qué está terminado y qué no para pasar a la siguiente fase.

VII. APORTE TEÓRICO Y SIGNIFICACION PRÁCTICA.

APORTE TEORICO.

El **reconocimiento óptico de caracteres (OCR)** es el proceso de convertir una imagen de texto en un formato de texto legible por máquina. Si, p. escanea un formulario o recibo, la computadora guarda la imagen escaneada como un archivo de imagen. No puede editar, buscar o contar palabras en archivos de imagen usando un editor de texto. Sin embargo, se puede usar OCR para convertir una

imagen en un documento de texto que contiene datos de texto. (¿Qué Es El OCR? - Explicación Del Reconocimiento Óptico de Caracteres - AWS, 2023)

El **reconocimiento de voz** es la capacidad de una máquina o programa para reconocer palabras y frases en el lenguaje hablado y convertirlas a un formato legible por máquina. El software de reconocimiento de voz para principiantes tiene un vocabulario limitado de palabras y frases y solo puede reconocerlas si habla con claridad. El software más sofisticado puede aceptar el habla natural. (González, 2020)

SIGNIFICACION PRÁCTICA.

Para destacar la función práctica del **OCR** se puede decir que: La mayoría de los procesos comerciales implican recibir información de medios impresos. Los formularios impresos, las facturas, los documentos legales escaneados y los contratos impresos son parte del proceso comercial. Este gran volumen de documentos requiere mucho tiempo y espacio para almacenar y administrar. (¿Qué Es El OCR? - Explicación Del Reconocimiento Óptico de Caracteres - AWS, 2023)

Brindar a los clientes el máximo poder de escucha en el **reconocimiento de voz** cuando necesitan ayuda es un elemento fundamental que impulsa el valor de la marca y es cada vez más estratégico, ya que la proximidad de una empresa a las necesidades del consumidor es fundamental para la ventaja competitiva. Pero implementar un centro de contacto, o incluso construir uno desde cero, significa tomar recursos financieros y desviar profesionales y expertos que podrían usarse para respaldar el crecimiento del negocio principal. El reconocimiento de voz garantiza el mismo tipo de servicio a un precio más bajo: gracias a la integración con tecnologías de inteligencia artificial, es posible introducir chatbots y asistentes virtuales que son capaces de comprender las solicitudes del interlocutor, filtrando preguntas que pueden gestionarse automáticamente y entregarse directamente sólo al servicio al cliente. Aquellos empleados, que realmente requieren intervención humana. (Usos Y Ventajas Del Reconocimiento Del Habla | Almawave, 2020)

IX.BIBLIOGRAFÍA.

- DE, P. (2023). PERFIL DE PROYECTO INEGRADOR - UNIANDES. Google.com. <https://sites.google.com/site/primerosistemas/proyecto-integra/perfil-de-proyecto-inegrador>
- Atlassian. (2016). Scrum: qué es, cómo funciona y por qué es excelente. Atlassian. <https://www.atlassian.com/es/agile/scrum>
- ¿Qué es el OCR? - Explicación del reconocimiento óptico de caracteres - AWS. (2023). Amazon Web Services, Inc. <https://aws.amazon.com/es/what-is/ocr/>
- González, A. (2020, May 19). Reconocimiento de voz: Que es, cómo funciona y programas que existen. Ayuda Ley Protección Datos; AyudaLeyProteccionDatos. [https://ayudaleyprotecciondatos.es/2020/05/19/reconocimiento-voz/#Que es el reconocimiento de voz](https://ayudaleyprotecciondatos.es/2020/05/19/reconocimiento-voz/#Que%20es%20el%20reconocimiento%20de%20voz)
- Usos y ventajas del reconocimiento del habla | Almawave. (2020). Almawave. <https://www.almawave.com/es/2020/03/05/usi-e-vantaggi-della-speech-recognition/>