



Universidad del Azuay

Facultad de Ciencias de la Administración

Carrera de Ingeniería de Sistemas y Telemática

**DESARROLLO DE UNA METODOLOGÍA PARA
GENERAR VÍNCULOS RELACIONADOS A
TEXTOS TRANSCRITOS PARA VIDEOS DE
YOUTUBE**

**Trabajo de titulación previo a la obtención del grado en
Ingeniero de Sistemas y Telemática**

Autores:

Pablo Daniel Martínez León

Director:

Ing. Marcos Patricio Orellana Cordero

Cuenca – Ecuador

2023

DEDICATORIA

Dedico este trabajo a ustedes, mis amados papitos, este logro no habría sido posible sin su amor, su apoyo incondicional y constante aliento a lo largo de toda mi vida. Por hacerme la persona que soy y que confío y crea en mí siempre. Su dedicación, sacrificio y valores han sido pilares fundamentales en mi formación y en mi día a día. Gracias por ser mis guías, mis consejeros y mis mayores admiradores. No hay palabras suficientes para expresar mi gratitud por todo lo que han hecho por mí. Los amo con todo mi corazón y gracias por siempre brindarme su cariño, por ser mi inspiración y por darme las raíces que me han sostenido y las alas que me han permitido volar. Esto es por y para ustedes.

A mis abuelos y abuelas, les agradezco por su sabiduría, consejos y amor incondicional. Sus palabras de aliento y su apoyo constante me han impulsado a superar obstáculos y a perseguir mis sueños desde pequeño. Su presencia en mi vida es un regalo invaluable y su legado será siempre una guía en mi camino. Gracias por darme ese amor único y enseñarme la importancia de ser una persona íntegra en todas las áreas de mi vida. No hay mejor ejemplo a seguir que ustedes.

A mi novia, por todo el cariño, el amor y el apoyo durante toda esta etapa. Gracias por acompañarme y estar ahí en cada momento, por todo lo que has hecho y sigues haciendo por mí y por nunca faltarme. Estoy muy agradecido de tenerte a mi lado y poder compartir estos momentos especiales contigo.

Finalmente a todos mis familiares y amigos incondicionales, por estar ahí en todos los momentos de mi camino.

A todos, espero que esta tesis sea solo el comienzo de muchas alegrías compartidas.

AGRADECIMIENTO

Agradezco a mi director de tesis, Marcos Orellana, por su inigualable experticia y dedicación a lo largo de este tiempo. Su guía y orientación fueron fundamentales para llevar a cabo este estudio con éxito.

De igual manera, quiero expresar mi profundo agradecimiento a todo el grupo de investigación del laboratorio L.I.D.I. Su invaluable ayuda y colaboración han sido imprescindibles en cada etapa de este proyecto.

Agradezco de todo corazón a mis padres por su apoyo incondicional, paciencia y por brindarme las fuerzas necesarias para cumplir esta importante etapa de mi vida. Su amor y aliento han sido mi motor y me han dado la confianza para enfrentar cualquier desafío.

También quiero extender mi agradecimiento a cada uno de los profesores que formaron parte de mi instrucción académica. Su generosidad al compartir sus conocimientos y su paciencia para responder a mis dudas han sido de gran valor para mi formación.

No puedo dejar de mencionar mi profunda gratitud hacia la Universidad del Azuay, por brindarme la oportunidad de realizar mis estudios y por fomentar un ambiente propicio para el aprendizaje y el crecimiento personal.

Por último, quiero agradecer a mis familiares, a mi novia y a mis amigos que fueron parte de esta increíble etapa de mi vida. Su compañía, amor, amistad y apoyo han hecho de este viaje una experiencia inolvidable.

A cada uno de ustedes, mi más sincero agradecimiento por ser parte de este camino y por contribuir de manera significativa a mi desarrollo personal y profesional.

ÍNDICE

Índice de contenido

DEDICATORIA.....	I
AGRADECIMIENTO.....	II
Índice de contenido.....	III
Índice de Tablas.....	V
Índice de Figuras.....	VI
Índice de anexos.....	VII
Resumen:.....	VIII
Abstract:.....	VIII
1. Introducción.....	1
2. Marco Teórico.....	4
3. Estado del Arte.....	8
4. Metodología.....	11
4.1. Extraer Transcripción.....	12
4.2. Aplicar Técnicas de Minería de texto.....	14
4.2.1. Extraer palabras claves.....	14
4.3 Extraer artículos de Wikipedia.....	17
4.4. Clasificar la relación de los temas de interés.....	18
4.5. Generar Mashup.....	20
5. Resultados.....	22
6. Evaluación.....	28
6.1 Goal Question Metric.....	28
6.2 Modelo de aceptación tecnológica (TAM):.....	30
6.2.1. Definición de variables.....	30
6.2.2. Hipótesis.....	31
6.2.3 Ejecución.....	32
6.3 Análisis e interpretación de los datos recolectados.....	33
6.3.1 Resultados de la Encuesta.....	33
6.4 Validación de las Hipótesis.....	34
6.5 Amenazas a la validez.....	35
7. Conclusiones.....	38
8. REFERENCIAS.....	39
ANEXOS.....	42
Anexo 1.....	42

Anexo 2..... 44

Índice de Tablas

Tabla 1. Componentes funcionales de Yake!	17
Tabla 2. Cuestionario para la medición de variables con TAM	32
Tabla 3. Estadística descriptiva de los resultados obtenidos para cada variable	33

Índice de Figuras

Figura 1. Metodología SPEM	12
Figura 2. Metodología SPEM técnicas de minería de texto	14
Figura 3. Página de inicio y campo de texto del enlace de YouTube.....	21
Figura 4. Gráfico de Resultados Temática Informática.....	23
Figura 5. Resultado Temática Informática dentro del Mashup	24
Figura 6. Gráfico de Resultados de la temática de matemáticas	25
Figura 7. Resultados de temática de matemáticas dentro del Mashup	26

Índice de anexos

Anexo 1. Cuestionario para la medición de variables con TAM.....	44
Anexo 2. Resultados del Cuestionario para la medición de variables con TAM.....	42

Resumen:

En la actualidad, las herramientas tecnológicas ayudan en tareas de aprendizaje. Existen diversas fuentes de información donde las personas se apoyan para complementar sus actividades académicas. Se busca contenido en la web, significativamente actualizado y fácil de entender, generalmente en forma de videos o textos cortos. Conforme las personas incrementan su aprendizaje, se enfrentan a términos y conceptos con los que no están familiarizados. Esto genera un proceso complejo de vincular palabras y frases de la transcripción del video con otras fuentes de información que puedan explicar los temas relacionados y que contribuyan a la comprensión del tema tratado. La presente investigación involucra el desarrollo de un método donde se integren diferentes fuentes de datos, a través de un software basado en técnicas de minería de texto. La misma que permitirá extraer el contenido textual de los videos e identificar los términos que se relacionan con el tema de estudio.

Palabras clave: Clasificación de texto, Mashup, Minería de texto, Palabras claves, Procesamiento de lenguaje natural

Abstract:

Nowadays, technological tools assist in learning tasks. There are various sources of information that people rely on to complement their academic activities. Content, significantly updated and easy to understand, usually in the form of videos or short texts is searched on the web. As people increase their learning, they are confronted with terms and concepts with which they are not familiar with. Therefore, a complex process is generated to link words and phrases from the video transcript with other sources of information that can explain the related topics and contribute to the understanding of the subject being addressed. This research involves the development of a method where different data sources are integrated through text mining techniques-based software. This method will allow extracting the textual content from videos and identifying the terms that are related to the topic of study.

Keywords: Keywords, Mashup, Natural Language Processing, Text Classification, Text Mining



Este certificado se encuentra en el repositorio digital de la Universidad del Azuay, para verificar su autenticidad escanee el código QR

Este certificado consta de: 1 página