



**Aplicación web para el análisis de contenido usando servicios cognitivos de  
Inteligencia Artificial en la plataforma web educativa Jovenesweb**

Betancourt Ludeña, Elvis Fernando y Panata Castillo, Edgar Santiago

Departamento de Ciencias de la Computación

Carrera de Ingeniería en Software

Trabajo de titulación, previo a la obtención del título de Ingeniero en Software

Ing, Velasco Sánchez, Diego Ismael

28 de julio de 2022

Latacunga

## Reporte de verificación de contenido



Trabajo de titulación sin caratula Betancourt Panata.pdf

Scanned on: 13:52 July 28, 2022 UTC



Overall Similarity Score



Results Found



Total Words in Text

Identical Words	589
Words with Minor Changes	155
Paraphrased Words	171
Omitted Words	0

A handwritten signature in blue ink, appearing to read "Diego Ismael Velasco Sánchez".

Ing. Diego Ismael Velasco Sánchez

CC: 0502603855



Website | Education | Businesses



Departamento de Ciencias de la Computación

Carrera de Ingeniería en Software

### Certificación

Certifico que el trabajo de titulación, “Aplicación web para el análisis de contenido usando servicios cognitivos de inteligencia artificial en la plataforma web educativa Jovenesweb” fue realizado por los señores **Betancourt Ludeña, Elvis Fernando** y **Panata Castillo, Edgar Santiago**; el mismo que cumple con los requisitos legales teóricos, científicos, técnicos y metodológicos establecidos por la Universidad de las Fuerzas Armadas ESPE, además fue revisado y analizado en su totalidad por la herramienta de prevención y/o verificación de similitud de contenidos; razón por la cual me permito acreditar y autorizar para que se lo sustente públicamente.

Latacunga, 28 de julio de 2022

A handwritten signature in black ink, which appears to read 'Diego Velasco S.', is written over a horizontal dotted line.

Ing. Velasco Sánchez, Diego Ismael  
C.C.: 0502603855



**ESPE**  
UNIVERSIDAD DE LAS FUERZAS ARMADAS  
INNOVACIÓN PARA LA EXCELENCIA

Departamento de Ciencias de la Computación

Carrera de Ingeniería en Software

Responsabilidad de Autoría

Nosotros, **Betancourt Ludeña, Elvis Fernando**, con cédula de ciudadanía 1104149719 y **Panata Castillo, Edgar Santiago**, con cédula de ciudadanía 1722403423, declaramos que el contenido, ideas y criterios del trabajo de titulación: **“Aplicación web para el análisis de contenido usando servicios cognitivos de inteligencia artificial en la plataforma web educativa Jovenesweb”**, es de nuestra autoría y responsabilidad, cumpliendo con los requisitos legales, teóricos, científicos, técnicos y metodológicos establecidos por la Universidad de las Fuerzas Armadas ESPE, respetando los derechos intelectuales de terceros y referenciado las citas bibliográficas.

Latacunga, 28 de julio de 2022

Betancourt Ludeña, Elvis Fernando  
C.C.: 1104149719

Panata Castillo, Edgar Santiago  
C.C.: 1722403423



**ESPE**  
UNIVERSIDAD DE LAS FUERZAS ARMADAS  
INNOVACIÓN PARA LA EXCELENCIA

Departamento de Ciencias de la Computación

Carrera de Ingeniería en Software

**Autorización de publicación**

Nosotros, **Betancourt Ludeña, Elvis Fernando**, con cédula de ciudadanía 1104149719 y **Panata Castillo, Edgar Santiago**, con cédula de ciudadanía 1722403423, autorizamos a la Universidad de las Fuerzas Armadas ESPE a publicar el trabajo de titulación: **"Aplicación web para el análisis de contenido usando servicios cognitivos de inteligencia artificial en la plataforma web educativa Jovenesweb"**, en el Repositorio Institucional, cuyo contenido, ideas y criterios son de nuestra responsabilidad.

Latacunga, 28 de julio de 2022

.....  
Betancourt Ludeña, Elvis Fernando  
C.C.: 1104149719

.....  
Panata Castillo, Edgar Santiago  
C.C.: 1722403423

### **Dedicatoria**

Dedico el presente proyecto a mis padres los cuales me han enseñado a perseverar y nunca rendirme para cumplir mis objetivos y de los cuales he aprendido muchas cosas.

A mis hermanas que siempre han estado apoyándome e incentivándome siempre a ser una mejor persona.

Finalmente, a mis maestros quienes han sabido guiarme durante mi formación profesional y me han motivado a no ser conformista y estar aprendiendo constantemente.

BETANCOURT LUDEÑA, ELVIS FERNANDO

### **Dedicatoria**

Con mucho afecto y cariño dedico este proyecto a mis padres, quienes con su ejemplo y dedicación me han enseñado a perseverar y esforzarme, además, han sido el motor fundamental durante mi etapa como estudiante universitario con su apoyo constante y paciencia. También de manera especial a mis hermanos que de una u otra manera han estado en todo momento con su mano amiga.

PANATA CASTILLO, EDGAR SANTIAGO

### **Agradecimiento**

Agradezco de manera muy especial a mis padres que siempre han estado apoyándome a conseguir mis metas y me han brindado todo el apoyo para poder realizar mis estudios de la mejor manera.

A mis hermanas que siempre han estado presente y me han motivado a estudiar y conseguir los objetivos que tengo planteados.

A mis compañeros con los que he compartido grandes momentos durante toda mi educación profesional en la universidad.

A mis profesores, quienes con paciencia y dedicación me han apoyado durante todo este proceso de formación, al Ing. Diego Velasco el cual nos ha apoyado con su conocimiento y tiempo para la elaboración del presenta trabajo de tesis.

Por último, a mi compañero de tesis Santiago P, con el que he compartido tiempo durante la elaboración del presente trabajo y además durante toda mi carrera universitaria.

BETANCOURT LUDEÑA, ELVIS FERNANDO



### **Agradecimiento**

De manera especial agradezco a mis padres por su apoyo contante en mi crecimiento profesional para conseguir mis metas, y siempre con su amabilidad y cariño me han incentivado en mi crecimiento personal.

A mis hermanos que han siempre han estado apoyando durante mi etapa de educativa e incentivándome a mejorar mis conocimientos.

Agradecer a mis compañeros con quienes compartí varias experiencias positivas que me ayudaron a crecer como persona.

A los docentes que me han transmitido su conocimiento y experiencia en las aulas de clase, siempre incentivándonos al crecimiento profesional y de manera especial al ing. Diego Velasco que con paciencia y sabiduría nos ha sabido guiar en el desarrollo del proyecto.

De igual manera agradecer a mi compañero Elvis B, quien con quien he compartido durante todo el periódico como estudiante.

PANATA CASTILLO, EDGAR SANTIAGO

**ÍNDICE DE CONTENIDO**

<b>Carátula .....</b>	<b>1</b>
<b>Reporte de verificación de contenido.....</b>	<b>2</b>
<b>Certificación .....</b>	<b>3</b>
<b>Responsabilidad de Autoría.....</b>	<b>4</b>
<b>Autorización de Publicación .....</b>	<b>5</b>
<b>Dedicatoria .....</b>	<b>6</b>
<b>Dedicatoria .....</b>	<b>7</b>
<b>Agradecimiento.....</b>	<b>8</b>
<b>Agradecimiento.....</b>	<b>9</b>
<b>Índice de contenido .....</b>	<b>10</b>
<b>Índice de figuras .....</b>	<b>15</b>
<b>Índice de tablas .....</b>	<b>17</b>
<b>Resumen.....</b>	<b>19</b>
<b>Abstract .....</b>	<b>20</b>
<b>Capítulo I: Introducción.....</b>	<b>21</b>
<b>Planteamiento del problema.....</b>	<b>21</b>
<b>Formulación del problema.....</b>	<b>23</b>
<b>Antecedentes.....</b>	<b>23</b>
<b>Justificación e importancia .....</b>	<b>25</b>
<b>Objetivos.....</b>	<b>26</b>

<i>Objetivo General</i> .....	26
<i>Objetivos Específicos</i> .....	26
<b>Hipótesis</b> .....	27
<b>Variables</b> .....	27
<i>Variable independiente</i> .....	27
<i>Variable dependiente</i> .....	28
<b>Capítulo II: Fundamentación conceptual y referencial</b> .....	29
<b>Introducción del capítulo</b> .....	29
<b>Antecedentes históricos</b> .....	29
<b>Antecedentes Conceptuales y Referenciales</b> .....	32
<i>Metodología ágil Scrum</i> .....	32
<i>Inteligencia artificial</i> .....	33
<i>Servicios de inteligencia artificial</i> .....	35
<i>Servicios cognitivos Microsoft Azure</i> .....	36
<i>Redes sociales</i> .....	37
<i>Moderación de contenido en las redes sociales</i> .....	38
<b>Base conceptual</b> .....	39
<i>Contenido web</i> .....	40
<i>Análisis del contenido</i> .....	40
<i>Contenido ofensivo</i> .....	41
<i>Sentimientos</i> .....	42

<i>Entidades</i> .....	42
<i>Base legal</i> .....	43
<b>Antecedentes Contextuales</b> .....	43
<b>Capitulo III: Desarrollo de la propuesta</b> .....	48
<b>Etapa 1: Inicio del proyecto</b> .....	48
<i>Visión del proyecto</i> .....	48
<i>Equipo Scrum</i> .....	48
<i>Definición de roles del Product Backlog</i> .....	49
<i>Product Backlog</i> .....	49
<i>Priorización del Backlog</i> .....	53
<i>Arquitectura del proyecto</i> .....	54
<i>Cronograma de entregas</i> .....	56
<b>Etapa 2: Planificación del Sprint</b> .....	57
<i>Primer Sprint Planning</i> .....	57
<i>Segundo Sprint Planning</i> .....	58
<i>Tercer Sprint Planning</i> .....	58
<i>Cuarto Sprint Planning</i> .....	59
<b>Etapa 3: Desarrollo del Sprint</b> .....	59
<i>Sprint 1</i> .....	60
<i>Sprint 2</i> .....	64
<i>Sprint 3</i> .....	70

<i>Sprint 4</i> .....	73
<b>Etapa 4: Revisión del Sprint</b> .....	76
<i>Entregable Sprint 1</i> .....	76
<i>Entregable Sprint 2</i> .....	80
<i>Entregable Sprint 3</i> .....	83
<i>Entregable Sprint 4</i> .....	86
<b>Etapa 5: Implementación</b> .....	88
<b>Capítulo IV: Validación de la propuesta</b> .....	92
Introducción del capítulo .....	92
Obtención de datos .....	92
<i>Datos recolectados</i> .....	93
Análisis de datos .....	95
Validación para el indicador 1 .....	95
Validación para el indicador 2 .....	97
Validación para el indicador 3 .....	98
Validación para el indicador 4 .....	100
Validación para el indicador 5 .....	101
Validación para el indicador 6 .....	102
Presentación de resultados del análisis de los datos .....	104
<b>Capítulo V: Conclusiones y recomendaciones</b> .....	105
Conclusiones .....	105

<b>Recomendaciones.....</b>	<b>106</b>
<b>Bibliografía .....</b>	<b>108</b>
<b>Anexos.....</b>	<b>114</b>

## ÍNDICE DE FIGURAS

<b>Figura 1</b> <i>Línea de Tiempo de la Inteligencia Artificial</i> .....	33
<b>Figura 2</b> <i>Ventajas de los servicios cognitivos</i> .....	37
<b>Figura 3</b> <i>Nivel 1 de la arquitectura del proyecto</i> .....	55
<b>Figura 4</b> <i>Nivel 1 de la arquitectura del proyecto</i> .....	56
<b>Figura 5</b> <i>Cronograma de entrega de los sprint a realizar</i> .....	57
<b>Figura 6</b> <i>Tablero de productividad del Sprint 1</i> .....	63
<b>Figura 7</b> <i>Tablero de productividad del Sprint 2</i> .....	70
<b>Figura 8</b> <i>Tablero de productividad del Sprint 3</i> .....	73
<b>Figura 9</b> <i>Tablero de productividad del Sprint 4</i> .....	76
<b>Figura 10</b> <i>Página principal de ContemModerator</i> .....	77
<b>Figura 11</b> <i>Diseño del modelo de la base de datos</i> .....	77
<b>Figura 12</b> <i>Página de creación de una cuenta</i> .....	79
<b>Figura 13</b> <i>Página de inicio de sesión</i> .....	80
<b>Figura 14</b> <i>Categorización de los registros ofensivos y no ofensivos</i> .....	81
<b>Figura 15</b> <i>Categorización de las entidades</i> .....	82
<b>Figura 16</b> <i>Documentacion de las APIs disponibles</i> .....	83
<b>Figura 17</b> <i>Generacion de reportes en PDF de cada identidad</i> .....	84
<b>Figura 18</b> <i>Generación de los reportes en Excel de cada entidad</i> .....	85
<b>Figura 19</b> <i>Generación de graficos estadísticos</i> .....	86
<b>Figura 20</b> <i>Porcentaje de contenido ofensivo y palabras más repetidas</i> .....	87
<b>Figura 21</b> <i>Gráfico estadístico de barras con las entidades mas repetidas</i> .....	87
<b>Figura 22</b> <i>Tablas dinámicas para el manejo de información</i> .....	88
<b>Figura 23</b> <i>Repositoria del proyecto</i> .....	89
<b>Figura 24</b> <i>Servidor que almacena el repositorio</i> .....	90

<b>Figura 25</b> <i>Gráfico de chi-cuadrado del indicador 1</i> .....	96
<b>Figura 26</b> <i>Gráfico de chi-cuadrado del indicador 2</i> .....	98
<b>Figura 27</b> <i>Gráfico de chi-cuadrado del indicador 3</i> .....	99
<b>Figura 28</b> <i>Gráfico de chi-cuadrado del indicador 4</i> .....	100
<b>Figura 29</b> <i>Gráfico de chi-cuadrado del indicador 5</i> .....	102
<b>Figura 30</b> <i>Gráfico de chi-cuadrado del indicador 6</i> .....	103



## ÍNDICE DE TABLAS

<b>Tabla 1</b> <i>Proveedores de servicios cognitivos Inteligencia Artificial</i> .....	35
<b>Tabla 2</b> <i>Historia de usuario 1</i> .....	49
<b>Tabla 3</b> <i>Historia de usuario 2</i> .....	50
<b>Tabla 4</b> <i>Historia de usuario 3</i> .....	50
<b>Tabla 5</b> <i>Historia de usuario 4</i> .....	50
<b>Tabla 6</b> <i>Historia de usuario 5</i> .....	50
<b>Tabla 7</b> <i>Historia de usuario 6</i> .....	51
<b>Tabla 8</b> <i>Historia de usuario 7</i> .....	51
<b>Tabla 9</b> <i>Historia de usuario 8</i> .....	51
<b>Tabla 10</b> <i>Historia de usuario 9</i> .....	51
<b>Tabla 11</b> <i>Historia de usuario 10</i> .....	52
<b>Tabla 12</b> <i>Historia de usuario 11</i> .....	52
<b>Tabla 13</b> <i>Historia de usuario 12</i> .....	52
<b>Tabla 14</b> <i>Historia de usuario 13</i> .....	52
<b>Tabla 15</b> <i>Historia de usuario 14</i> .....	53
<b>Tabla 16</b> <i>Product backlog priorizado</i> .....	53
<b>Tabla 17</b> <i>Planning del primer sprint</i> .....	57
<b>Tabla 18</b> <i>Planning del segundo sprint</i> .....	58
<b>Tabla 19</b> <i>Planning del tercer sprint</i> .....	58
<b>Tabla 20</b> <i>Planning del cuarto sprint</i> .....	59
<b>Tabla 21</b> <i>Criterios de aceptación de la historia de usuario 1</i> .....	60
<b>Tabla 22</b> <i>Criterios de aceptación de la historia de usuario 2</i> .....	61
<b>Tabla 23</b> <i>Criterios de aceptación de la historio de usuario 3</i> .....	61
<b>Tabla 24</b> <i>Criterios de aceptación de la historia de usuario 4</i> .....	62

<b>Tabla 25</b> <i>Criterios de aceptación de la historia de usuario 5</i> .....	64
<b>Tabla 26</b> <i>Criterios de aceptación de la historia de usuario 6</i> .....	65
<b>Tabla 27</b> <i>Criterios de aceptación de la historia de usuario 7</i> .....	67
<b>Tabla 28</b> <i>Criterios de aceptación de la historia usuario 8</i> .....	69
<b>Tabla 29</b> <i>Criterios de aceptación de la historia de usuario 9</i> .....	70
<b>Tabla 30</b> <i>Criterios de aceptación de la historia de usuario 10</i> .....	71
<b>Tabla 31</b> <i>Criterios de aceptación de la historia de usuario 11</i> .....	72
<b>Tabla 32</b> <i>Criterios de aceptación de la historia de usuario 12</i> .....	73
<b>Tabla 33</b> <i>Criterios de aceptación de la historia de usuario 13</i> .....	74
<b>Tabla 34</b> <i>Criterios de aceptación de la historia de usuario 14</i> .....	75
<b>Tabla 35</b> <i>Frecuencia esperada y frecuencia observada del indicador 1</i> .....	93
<b>Tabla 36</b> <i>Frecuencia esperada y frecuencia observada del indicador 2</i> .....	93
<b>Tabla 37</b> <i>Frecuencia esperada y frecuencia observada del indicador 3</i> .....	94
<b>Tabla 38</b> <i>Frecuencia esperada y frecuencia observada del indicador 4</i> .....	94
<b>Tabla 39</b> <i>Frecuencia esperada y frecuencia observada del indicador 5</i> .....	94
<b>Tabla 40</b> <i>Frecuencia esperada y frecuencia observada del indicador 6</i> .....	94
<b>Tabla 41</b> <i>Calculo de bondad de ajuste de chi-cuadrado del indicador 1</i> .....	96
<b>Tabla 42</b> <i>Calculo de bondad de ajuste de chi-cuadrado del indicador 2</i> .....	97
<b>Tabla 43</b> <i>Calculo de bondad de ajuste de chi-cuadrado del indicador 3</i> .....	99
<b>Tabla 44</b> <i>Calculo de bondad de ajuste de chi-cuadrado del indicador 4</i> .....	100
<b>Tabla 45</b> <i>Calculo de bondad de ajuste de chi-cuadrado del indicador 5</i> .....	101
<b>Tabla 46</b> <i>Calculo de bondad de ajuste de chi-cuadrado del indicador 6</i> .....	103

## RESUMEN

El presente proyecto está orientado al desarrollo de una aplicación web para optimizar el análisis de contenido que publican los usuarios de la plataforma Jovenesweb, la misma que cuenta con más de un millón de usuarios y brinda el servicio educativo, dado que el análisis de forma manual se caracteriza por ser costoso y demoroso, además, trae consecuencias negativas hacia la persona que lo realiza como: la normalización de contenido violento, problemas emocionales. También se debe considerar que el análisis pierde objetividad al intervenir la cultura, creencias y educación de quien realiza el análisis del contenido, para el desarrollo se ha utilizado un marco de trabajo interactivo incremental Scrum por su adaptabilidad y flexibilidad frente a los cambios, además, se configuro e Íntegro con los servicios cognitivos de Inteligencia Artificial de Microsoft Azure, de esta manera se realiza graficas estadísticas y tablas dinámicas con la información generada del análisis de contenido ofensivo, entidades, y los sentimientos principales (positivo, negativo y neutral), la validación del proyecto está bajo un enfoque del cumplimiento de los indicadores con el análisis estadístico del chi-cuadrado por bondad de ajuste, es importante indicar que para el análisis se debe normalizar los textos y existe limitaciones del lenguaje en el análisis como el uso de las reglas gramáticas y las abreviaturas de las de palabras y frases.

*Palabras claves:* Aplicativo web, Análisis de contenido, Servicios cognitivos, Inteligencia Artificial

## ABSTRACT

This project is oriented to the development of a web application to optimize the analysis of content published by users of the Jovenesweb platform, which has more than one million users and provides educational service, since manual analysis is characterized by being costly and time consuming, in addition, it brings negative consequences to the person who performs it as: the normalization of violent content, emotional problems. It should also be considered that the analysis loses objectivity by intervening the culture, beliefs and education of the person who performs the content analysis, for the development an incremental interactive framework Scrum has been used for its adaptability and flexibility to changes, in addition, it was configured and integrated with the cognitive services of Artificial Intelligence of Microsoft Azure, in this way statistical graphs and pivot tables with the information generated from the analysis of offensive content are performed, entities, and the main feelings (positive, negative and neutral), the validation of the project is under an enforce of the fulfillment of the indicators with the statistical analysis of the chi-square by goodness of fit, it is important to indicate that for the analysis the texts must be normalized and there are limitations of language in the analysis as the use of grammar rules and abbreviations of words and phrases.

*Keywords:* Web application, Content analysis, Cognitive services, Artificial Intelligence

## Capítulo I

### Introducción

#### Planteamiento del problema

El creciente uso de las redes sociales ha provocado el envío de millones de mensajes y publicaciones creando contenido en sitios populares como Instagram, Twitter, Facebook y YouTube. Parte del contenido pueden llegar a ser molestos u ofensivos, los mismos afectan de manera especial a niños, adolescentes y adultos jóvenes, también llegan a tener gran repercusión provocando desestabilidad emocional o problemas en su salud mental, además tienden a incentivar el odio en contra de grupos sociales dado su grupo étnico, orientación sexual o creencias religiosas (Balaguer, 2019).

La moderación de contenido en las redes sociales más populares en un alto porcentaje es realizada por medio del talento humano, en el artículo “El horroroso trabajo del moderador de contenido online” del periódico El MUNDO del 11 de mayo del 2017 recolecta testimonios acerca de las consecuencias físicas y mentales que tuvieron después de abandonar el trabajo; tales como: dolores de cabeza, insomnio, estereotipos de grupos sociales, normalización de violencia y problemas mentales (Bruno, 2017).

Además, cualquier otro sitio web donde el usuario pueda expresar una opinión es de gran valor las mismas se pueden clasificar en opiniones positivas, negativas o neutrales; esta información clasificada por sentimiento ayuda en distintos eventos sociales o políticos, por ejemplo en las campañas políticas se puede conocer la aceptación de cada personaje político que se encuentre participando o ejerza influencia, en el estudio realizado por Gómez, Jaimes, Hidalgo y Luján (2018) sobre la Influencia de redes sociales en el análisis de sentimiento aplicado a la situación política en Ecuador a través de un análisis de sentimientos sobre el contenido de usuarios.

En la red social Twitter pudieron identificar la tendencia de apoyo que tenía en ese entonces el presidente. En otro ámbito el estudio que realizó Arboleda (2021) sobre el Análisis Comparativo Del Tratamiento De La Información Y La Opinión Pública Sobre El Paro De octubre 2019. Caso De Estudio: El Universo Y Wambra Medio Digital donde realizó un análisis de las publicaciones realizadas en ambos medios y las opiniones de los usuarios en cada uno de ellos se pudo identificar a los ciudadanos con más interés en los titulares que tengan palabras como muertos, heridos y en general todo lo relacionado a violencia en el ámbito del paro de octubre 2019, donde se realizó el estudio, por ende nos darnos cuenta de la necesidad de un aplicativo que permita analizar el contenido que aportan los usuarios tanto en las redes sociales como en sitios web en donde el usuario puede generar contenido.

Jovenesweb es una plataforma web y móvil gratuita para la preparación del examen de Acceso a la Educación Superior que evalúa la Secretaría de Educación Superior Ciencia Tecnología e Innovación (SENESCYT) es el ente encargado de regular el acceso a la educación superior, el cual basado en el artículo 81 de la Ley Orgánica de Educación Superior y en el segundo inciso del artículo 356 de la Constitución de la República del Ecuador se menciona que el ingreso a las instituciones de educación superior públicas se regula a través del Sistema de Nivelación y Admisión, para todos los y las aspirantes. El sistema se rige por los principios de méritos, igualdad de oportunidades y libertad de elección de carrera o carreras e institución. Jovenesweb cuenta con más de 1 millón de usuarios y permite a sus usuarios interactuar entre ellos, a través de comentarios y soluciones de las diferentes preguntas que se encuentran disponibles en la plataforma, es aquí donde se generan problemas para poder filtrar el contenido que los usuarios envían, dado que muchas veces suelen enviar textos ofensivos, negativos o que no tiene relación con el tema que están estudiando, esto conlleva un gasto.

En el caso de continuar con la moderación de contenido de forma manual, es posible que los tiempos de revisión se sigan extendiendo, además, es muy difícil revisar todo el contenido que aportan los usuarios, por lo que se seguirán acumulando hasta llegar a un punto donde sería necesario contratar a más personas para que puedan realizar la revisión, esto conlleva más gastos económicos, más tiempo y a pesar de esto no se garantiza la revisión al 100% de todos los mensajes.

Por tal razón el trabajo de investigación presentado en este proyecto tiene como finalidad el desarrollo de un sistema software que analiza el contenido generado por los usuarios a través de servicios cognitivos de inteligencia artificial optimizando el tiempo de revisión del contenido en la plataforma educativa Jovenesweb.

### **Formulación del problema**

En consecuencia, a la problemática descrita se obtiene el siguiente problema:

¿Cómo optimizar el proceso del análisis de contenido que publican los usuarios de la plataforma web educativa Jovenesweb?

### **Antecedentes**

La inteligencia cognitiva forma parte de la inteligencia artificial, comprende principalmente en imitar la inteligencia cognitiva del ser humano de observar y analizar lo observado, pero en el análisis de grandes cantidades de datos la tecnología no produce errores por cansancio o fastidio de tareas repetitivas, generando herramientas con capacidad de análisis, además, orientada a la interacción humana la misma que debe ser de forma natural y cotidiana, las principales ramas de la inteligencia cognitiva son el escucha y habla, procesamiento de lenguaje natural, comprensión de emociones y sentimientos, reconocimientos de entidades. Con ayuda de estos permite a las aplicaciones humanizar el proceso de comunicación entre las personas y un ambiente digital (García P. , 2021).

Existen diversas compañías tecnológicas que ofrecen servicios cognitivos de Inteligencia Artificial con un conjunto de redes neuronales, uno de ellos es Microsoft las cuales permiten a los desarrolladores usar sistemas de aprendizaje automático sin la necesidad de crear ni implementar modelos propios. Estos servicios cognitivos permiten integrar modelos de Inteligencia Artificial para la toma de decisiones, para el procesamiento del lenguaje natural, reconocimiento de sonidos de conversaciones y visión de computadora para analizar e identificar contenido en imágenes y videos.

En el año 2017 los autores Estevan Gómez-Torres, Roger Jaimes, Orlando Hidalgo y Sergio Luján-Mora realizaron un estudio para conocer la influencia de las redes sociales a la situación política en Ecuador a través del análisis de sentimientos en textos obtenidos en la plataforma Twitter, para analizar los textos los autores usaron una metodología Stanford NLP equivalente a modelos de regresión logística multiclase, como resultado obtuvieron que los datos analizados en Twitter no coincidieron con los resultados de las elecciones de Ecuador en 2017 y que los algoritmos de aprendizaje prometen soluciones en el área del procesamiento del lenguaje natural, así como la necesidad de generar nuevas técnicas para análisis de texto.

Diego Enríquez (2018) acota que en su investigación realizada se encontraron algunos tipos de complejidades una de ellas es al realizar el análisis de sentimiento de un personaje público del Ecuador, el idioma español es muy complejo a la hora de interpretarlo debido que se encuentran sesgos que complica mucho el análisis, otro de las limitantes es que en el país la red social de Twitter no es muy utilizada.



Álvaro Cardona (2020) integró los servicios cognitivos de Microsoft Azure para desarrollar un sistema de servicio al cliente con reconocimiento de intención y análisis de voz y texto por medio de Microsoft Dynamics Customer Service donde concluyó que esta solución ofrece una posibilidad de negocio debido a una mayor satisfacción de cliente por una gran fluidez en la atención de incidencias y el ahorro en recursos humanos gracias a la automatización.

Según Daniel Cervantes (2020) para realizar un correcto análisis de sentimientos en textos usando Inteligencia Artificial es necesario realizar un preprocesamiento de la información para que los datos resultantes puedan estar lo más limpio y normalizado posible, a través de varias técnicas que pueden ser: tratamiento de la duplicidad de caracteres, normalización de risas (jajaja por ja, jejeje por ja), eliminación de enlaces, normalización de jerga y normalización de mayúsculas y minúsculas.

### **Justificación e importancia**

La mayoría de las plataformas, redes sociales, foros, chats y plataformas educativas nacen sin límites de expresión. Dentro de las paredes virtuales se puede compartir o publicar lo que deseas, sin embargo, esto se puede convertir en mensajes de odio, racismo, palabras ofensivas, creando un ambiente de violencia.

Entonces nace la necesidad de la moderación del contenido en las plataformas en las que los usuarios pueden interactuar con otros usuarios, en un principio usando filtros humanos. con el avance del desarrollo de algoritmos de aprendizaje en la inteligencia artificial (IA) se puede automatizar la revisión de los mensajes publicados con el adecuado entrenamiento en el análisis de texto, en el caso de análisis de imágenes y videos aun necesitan pasar por un filtro humano.

El presente proyecto por medio de una aplicación web contribuye a optimizar el tiempo y calidad del análisis de los mensajes publicados por los usuarios en la plataforma educativa Jovenesweb, con ayuda de los servicios cognitivos de inteligencia artificial, en el cual se identificarán palabras ofensivas, sentimientos que emiten los mensajes y entidades que corresponden a información como: correo electrónico, número telefónico, nombres de personas y entidades generales como: fechas, números, lugares e información personal.

La plataforma educativa Jovenesweb será beneficiada con el presente proyecto al lograr analizar de manera más efectiva los contenidos que son publicados, ayudando de esta manera automatizar procesos de revisión y con los datos representados en reportes y gráficas estadísticas ayudará a tomar decisiones basadas en información del comportamiento de los usuarios y no empirismo. Además, el aplicativo web al recibir la información de los mensajes por medio de una API se integra con facilidad a cualquier software, permitiendo que los desarrolladores externos solo se enfoquen en la lógica de negocios y de manera sencilla conocería el comportamiento de los usuarios.

## **Objetivos**

### ***Objetivo General***

- Desarrollar una aplicación web para optimizar el análisis de contenido que publican los usuarios de la plataforma web educativa Jovenesweb, usando servicios cognitivos de Inteligencia Artificial.

### ***Objetivos Específicos***

- Elaborar el marco teórico con los conceptos fundamentales necesarios para el desarrollo de una aplicación web que analice el contenido de la plataforma educativa Jovenesweb a través del uso de servicios cognitivos de Microsoft Azure.

- Configurar e integrar los servicios cognitivos a través de la API Rest de Microsoft Azure
- Desarrollar el aplicativo web aplicando el marco de trabajo SCRUM
- Implementar el aplicativo web en la plataforma educativa Jovenesweb ubicada en la ciudad de Loja, provincia de Loja.
- Validar el aplicativo web mediante pruebas de funcionalidad y rendimiento.

## **Hipótesis**

¿Si se desarrolla una aplicación web para el análisis de contenido usando servicios cognitivos de inteligencia artificial entonces se optimizan los tiempo y calidad de revisión en la plataforma web educativa Jovenesweb?

## **Variables**

### ***Variable independiente***

Aplicación web con servicios cognitivos de Inteligencia Artificial (IA) para el análisis de contenidos

### **Conceptualización de la variable independiente**

El aplicativo web es un software que da la facilidad de tener acceso desde cualquier lugar y dispositivo que esté conectado a internet y tenga un browser, este aplicativo web será realizado con el uso de servicios cognitivos, integra el análisis de texto, para la identificación contenido ofensivo, sentimientos y entidades, las entidades conlleva a toda la información que se obtenga de un texto, tal como ciudades, información personal (número de teléfono, nombres, correo, edad ), nombres de ciudades, nombres de personajes públicos y objetos o sustancias comunes, para este proyecto se definirá como contenido a las preguntas, explicaciones y comentarios aportados por los usuarios en la plataforma Jovenesweb.

Durante la ejecución del aplicativo web, presentara los datos del contenido transformado en información por medio de gráficos estadísticos y tablas informativas que permita tomar decisiones a los administradores de la plataforma educativa Jovenesweb.

***Variable dependiente***

- Optimiza el tiempo de revisión del contenido de la plataforma educativa Jovenesweb
- Optimiza la calidad en la revisión del contenido de la plataforma educativa Jovenesweb

**Indicadores**

- Porcentaje de sentimientos positivos, negativos y neutrales de los textos analizados
- Porcentaje de contenido ofensivo detectado
- Tiempo en el análisis de contenido
- Porcentaje del contenido analizado
- Cantidad de entidades detectadas y clasificadas
- Cantidad de contenido con sentimientos identificados y clasificado (positivo, negativo y neutro)

## Capítulo II

### Fundamentación conceptual y referencial

#### Introducción del capítulo

En el presente capítulo se realiza una recopilación de información de antecedentes históricos, conceptuales, referenciales y contextuales pertenecientes a los distintos métodos de moderación de contenido de plataformas web así como modelos arquitectónicos para su implementación.

Se abordan los distintos servicios de inteligencia artificial que proveen las empresas tecnológicas para el análisis de contenido necesarios para llevar a cabo el proyecto de manera eficiente y acorde a las necesidades planeadas por Jovenesweb.

#### Antecedentes históricos

Principios de internet, finales de la década de los 60 aparecería otro proyecto que cambiaría la manera en que nos comunicaríamos en el futuro. Dicho proyecto era nombrado ARPANET, fue la primera red interconectada de ordenadores. Sus inicios fueron los de unir las universidades de UCLA y Stanford en los Estados Unidos de América por medio de conmutadores de paquetes llamados procesadores de mensaje de interfaz, conjunto de circuitos telefónicos y computadores host, de una línea telefónica. Esta red de ordenadores fue desarrollada por la agencia de investigación avanzada del Departamento de Defensa de los Estados Unidos de América (actualmente llamado DARPA2) (Barcelona, 2018).

Teniendo en cuenta que una de las razones principales para el desarrollo de ARPANET fue de compartir recursos con los diferentes usuarios en tiempo real, más tarde, en los principios de la década de los 70 los científicos de ARPANET conectaron 40 puntos en diferentes localidades.

En un comienzo la comunicación de los ordenadores usaban NCP como protocolo de comunicación de extremo a extremo, sin embargo, este protocolo tenía la limitante al direccionar redes y maquinas, entonces, el equipo de ARPANET decide desarrollar una nueva versión del protocolo que pudiera soportar un entorno de red de arquitectura abierta, este nuevo protocolo se llamaría Protocolo de control de transmisión/Protocolo de internet (TCP/IP), permitiendo el desarrollo generalizado de LANS, PC y estaciones de trabajo, dando inicio a lo que hoy conocemos como internet (Leiner, Cerf, Kahn, & Wolff, s.f.).

En el CERN (Organización Europea para la investigación nuclear), en un principio se consideró y desarrollo la Web bajo la idea de automatizar el intercambio de la información entre los científicos de las diferentes universidades, haciendo uso de las tecnologías de la evolución de aquel entonces, en 1990 se formaliza la propuesta que englobaba conceptos principales. El documento contenía proyecto en hipertexto cuyo nombre era “WorldWideWeb” que se podía visualizar en los navegadores (CERN, 2019).

En 1990, se pone en funcionamiento el primer servidor y navegador web, la primera dirección de un sitio web en el mundo y de un servidor fue <http://info.cern.ch/hypertext/WWW/TheProject.html> el mismo que contenía información sobre el proyecto.

En los inicios de 1993, la Universidad de Illinois publica la primera versión de su navegador Mosaic desarrollado para ejecutar en un sistema operativo de X Windows, era amigables con los usuarios y ofrecía un sistema de ventanas, luego de unos meses se navegadores para computadoras personales y Macintosh.

Esto ayudo una expansión exponencial de la WWW, el CERN puso a disposición toda la documentación y código fuente a la comunidad y sin regalías convirtiéndolo en software libre (CERN, 2019).

Con el creciente agigantado de la internet y la WWW empezaron a crearse las primeras redes sociales y los primeros blogs. Los primeros blogs eran como diarios personales donde un usuario comenta cosas que le intrigan o aspectos y vivencias del su día a día. Los otros usuarios entraban y leían e incluso comentaban lo que decía el Blogger, Por otro lado, unas de las primeras redes sociales que tuvo mucho éxito y que tenía muchos usuarios fue MySpace. Era la primera gran comunidad virtual y pasó a ser una de las primeras comunidades virtuales donde la gente le dedicaba muchos minutos al día, haciendo que entre el 2005-2008 fuera la web más visitada en el mundo (Britannica, 2022). Pero en el camino de MySpace se cruzó un proyecto que redefiniría el concepto de comunidad virtual y lo evolucionaría en lo que hoy en día conocemos como red virtual.

Facebook inicialmente fue pensada para unir en una comunidad a los estudiantes de diferentes redes sociales. Resulta curioso que ARPANET también inicialmente fue pensada para unir diferentes universidades. Como sucedió con ARPANET o Internet, Facebook en septiembre del 2006 se abre a todo tipo de usuarios. Durante estos años las comunidades de internet son relativamente nuevas y abiertas al público en masas y los problemas empiezan aparecer, con el uso de cuentas falsas y lograr de cierta manera estar de manera anónima las personas empiezan a realizar publicaciones que afectan a la integridad y al crecimiento de las comunidades de las redes sociales, aparecen los moderadores humanos y sistemas de reconocimiento de determinadas palabras ofensivas o que promueven el odio.

En la actualidad. - La moderación automática se presenta como la mejor solución, ya que se trata de un software que trabaja 24x7, sin la necesidad de mantener un equipo humano junto a esa cobertura.

Al responder de manera automática, se obtiene una respuesta inmediata, permitiendo una excelente experiencia de usuario (en el caso de la pre-moderación), y eliminando la ventana de exposición propia de un modelo de post moderación.

Tratándose de sistemas desarrollados en arquitecturas altamente escalables, es también posible moderar picos de tráfico, sin tener que reorganizar equipos de moderadores con sobre turnos y pudiendo responder en tiempo real a los requisitos de moderación.

Finalmente, estas características redundan en una significativa reducción de los costos destinados a la moderación.

## **Antecedentes Conceptuales y Referenciales**

### ***Metodología ágil Scrum***

En el estudio realizado por Ramírez, Salgado, Ramírez, Manrique, Osuna y Rosales (2019) se menciona que la metodología SCRUM es una:

Una metodología que integra las mejores prácticas y el trabajo en equipo de colaboración para lograr mejores resultados a través de la colaboración de todos los diferentes roles que tiene Scrum, además muestra entregables parciales del proyecto en cada sprint ejecutado. Estos entregables se priorizan según su contribución al proyecto y son evaluados por los usuarios finales del proyecto, el Product Owner y el equipo de desarrollo.



Scrum se recomienda para proyectos en entornos complejos donde los requisitos cambian y también es importante para lograr resultados rápidos donde la innovación, la flexibilidad y la productividad son esenciales y quieren trabajar en un ambiente ágil.

### ***Inteligencia artificial***

Rouhiainen (2018) define a la Inteligencia Artificial como “la capacidad de las máquinas para usar algoritmos, aprender de los datos y utilizar lo aprendido en la toma de decisiones tal y como lo haría un ser humano”, a lo largo de los años la Inteligencia Artificial ha ido evolucionando por lo que a continuación se presenta una línea de tiempo que describe los avances que ha tenido la Inteligencia Artificial.

Entonces se considera a la Inteligencia Artificial como una rama de la informática que pretende responder afirmativamente a la pregunta que planteo Alan Turing, para lograr similar a la inteligencia de un ser humano con la inteligencia de las maquinas, en la actualidad el uso de la Inteligencia Artificial tiene muchos debates y cuestionamientos, algunos consideran que en el futuro puede traer más problemas que soluciones (Buitin, 2020).

### **Figura 1**

*Línea de Tiempo de la Inteligencia Artificial*



*Nota.* El gráfico corresponde a la evolución de la Inteligencia Artificial desde 1936 hasta la actualidad

En la Figura 1. se puede evidenciar como Alang Turing fue el pionero en la Inteligencia Artificial a través de su máquina que fue capaz de autogenerar cálculos sin necesidad de la intervención humana, también inspiró el conocido Test de Turing el cual permite evaluar si una computadora tiene el comportamiento inteligente como el de un humano, luego en 1943 Warren McCulloch y Walter Pitts proponen un modelo de red neuronal que utiliza conceptos como la fisiología básica y propusieron la actualización de neuronas conocida como Aprendizaje Hebbiano o de Hebb, en el año de 1955 los autores John McCarthy, Marvin Minsky y Claude Shannon acuñaron el término de Inteligencia Artificial para referirse al ingenio de las máquinas para realizar cálculos, en 1982 se creó el primer sistema experto comercial que representó un ahorro de 40 millones de dólares anuales, luego de varios avances en el año 2000 se otorga premio al primer programa con personalidad propia (Martínez Ortega & Medina Chicaiza, 2020)

Finalmente, en el año 2020 se anuncia a GPT-3 un modelo de inteligencia artificial entrenado con alrededor de 175000 millones de parámetros obtenidos de fuentes de internet y capaz de realizar varias acciones como mantener conversaciones con una persona, transformar ecuaciones del lenguaje humano a LaTeX, generación de código de programación con solo darle a conocer un par de instrucciones y muchas más aplicaciones que se encuentran en desarrollo.

### ***Servicios de inteligencia artificial***

Actualmente Amazon Web Services (AWS), Google, IBM y Microsoft Azure son las empresas que lideran el mercado de servicios de Inteligencia Artificial los cuales permiten a los desarrolladores el uso de estos para que puedan integrarlos en su propia plataforma y según sus propias necesidades. Existe una gran variedad de servicios en distintos campos los cuales se resumen en la Tabla 1.

**Tabla 1**

*Proveedores de servicios cognitivos Inteligencia Artificial*

<b>Proveedor</b>	<b>Imágenes</b>	<b>Videos</b>	<b>Texto</b>
Amazon Web Services	Celebridades Escena Humor Identificación Lectura Moderación Objetos Rostros	Actividad Celebridades Escena Humor Identificación Lectura Moderación Objetos	Entidades Intención Key Phrase Lenguaje Sentimiento Traducción Tópicos
Google	Celebridades Colores Escena Humor Lectura Logos Lugares Moderación Objetos Reversa Rostros	Escena Humor Moderación Objetos Transcripción	Entidades Filtro Intención Key Phrase Lenguaje Sentimiento Sintaxis Traducción Tópicos
IBM	Comidas Humor Lectura Moderación Objetos Rostros		Entidades Intención Key Phrase Lenguaje Metadata Personalidad Relaciones

Proveedor	Imágenes	Videos	Texto
			Autocompletar
	Celebridades	Escena	Entidades Filtro
Microsoft	Colores Escena	Celebridades	Intención Key
	Humor Identificación	Identificación Humor	Phrase Lenguaje
Azure	Lectura Lugares	Lectura Marcas	Ortografía
	Moderación Objetos	Moderación Objetos	Relaciones
	Rostros	Traducción	Sentimiento
		Transcripción	Sintaxis Traducción
			Tópicos

De la Tabla 1, se destaca que Google lidera los servicios de inteligencia artificial de imágenes a diferencia de los demás proveedores, en cuanto a los servicios de video se puede evidenciar como todos ofrecen este servicio a excepción de IBM, en los servicios de voz todos los proveedores lo ofrecen por igual, finalmente se puede destacar que Microsoft Azure lidera los servicios de texto ya que ofrece más alternativas que sus demás competidores.

### ***Servicios cognitivos Microsoft Azure***

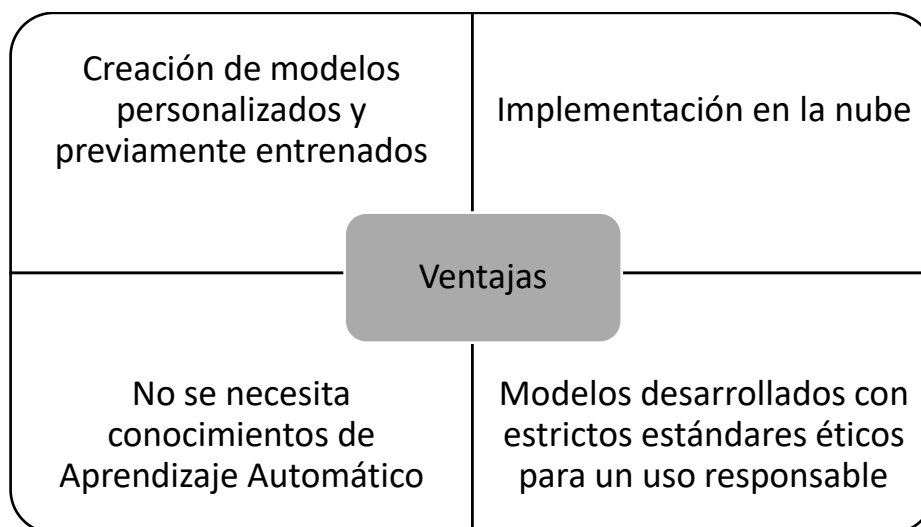
Microsoft Azure permite crear soluciones en la nube de manera segura y en diferentes entornos, actualmente ofrecen más de 200 productos de los cuales destacan productos de Inteligencia artificial y aprendizaje automático, cálculo, contenedores, soluciones híbridas y multinube e Internet de las cosas (IoT) (Servicios de informática en la nube | Microsoft Azure., 2021)

Entre los diferentes productos que ofrece Microsoft Azure se encuentran los Servicios Cognitivos, los cuales permiten a los desarrolladores usar la Inteligencia Artificial a través de una API sin la necesidad de tener experiencia en aprendizaje automático, con esto los desarrolladores pueden desarrollar sistemas inteligentes que le permitan ver, escuchar, hablar, buscar, comprender y acelerar la toma de decisiones en sus aplicaciones (Cognitive Services: API para desarrolladores de inteligencia artificial | Microsoft Azure, 2021)

En el siguiente gráfico podemos ver las ventajas de usar los servicios cognitivos de Microsoft Azure.

## Figura 2

### *Ventajas de los servicios cognitivos*



*Nota.* Esta figura describe las ventajas de usar los servicios cognitivos

En la Figura 2 podemos evidenciar como los servicios cognitivos nos permiten implementar modelos de inteligencia artificial ya entrenados de manera fácil y adaptable a las necesidades de cada aplicación.

### **Redes sociales**

Las redes sociales “son entornos digitales que permiten a los usuarios diseñar y visualizar sus perfiles, a partir de las acciones de compartir textos, imágenes, aplicaciones y otros contenidos, que promueven establecer una relación de información/comunicación con otros usuarios de la red” (Lara-Navarra, López-Borrull, Sánchez-Navarro, & Yànez, 2018).

De acuerdo con Maldonado, García y Sampedro (2019) las redes sociales “han tenido un crecimiento constante desde su creación, rompiendo barreras geográficas y paradigmas de comunicación que se han ido remplazando por nuevos sistemas virtuales que permiten la interacción casi de forma instantánea”.

Para estos autores la teoría de usos y gratificaciones analiza la conducta que tiene el sujeto con estos medios, por lo que se considera al destinatario como un ente activo y que participa de los medios de comunicación y afirma que los medios masivos no influyen en el espectador, sino que este modifica los contenidos que según sus necesidades.

El estudio realizado por López, Vizcaíno y Montes (2020) sobre los comentarios en las redes sociales: Los jóvenes asumen que las quejas o elogios sobre un determinado servicio, producto, evento o similar, es decir, lo expresado en el contenido publicado por parte de los usuarios puede influir en la toma de decisiones de otros usuarios. Los jóvenes también expresaron que existe un temor hacia el juicio público que pueden obtener en las redes sociales evitando en algunos momentos publicar sus verdaderos criterios.

### ***Moderación de contenido en las redes sociales***

Para Calcano (2021), encontró lo siguiente respecto al sistema de moderación de contenido de Facebook:

Un claro ejemplo es Facebook por medio de su sistema de administración de contenido logra revisar 2 millones de publicaciones generadas por los usuarios para determinar si cumplen o no los estándares y reglas de la comunidad, si el contenido analizado no cumple con las normas se elimina de la red social, algunos momento las publicaciones pueden estar sujetas a las subjetividad de la persona que la revisa quien puede considerar que cierto contenido es o no ofensivo.

En este sistema global se encuentran alrededor de 30 000 personas en diez oficinas que se encuentran alrededor del mundo especialmente en países asiáticos, se debe recordar que este sistema no lo maneja directamente Facebook sino es subcontratado por terceros, además incluye expertos de libertad de expresión, varios talleres dados por académicos y apoyo por parte de la Inteligencia Artificial.

También Calcano (2021), menciona lo siguiente sobre los sistemas de moderación de contenido de Twitter y YouTube: en estas dos empresas tiene un sistema similar a Facebook hoy en día conocido como meta, se debe considerar que la moderación de contenido de redes sociales con interacción a nivel mundial analiza más allá de contenido ofensivo, sino que también pueden ser falsedades que se propagan a través de comentarios o publicaciones hechas por los usuarios en las redes sociales, y para analizar este tipo de contenidos se requiere una conexión con diversos sitios de internet provocando un ejercicio muy complicado para la Inteligencia Artificial de la actualidad. Para Sylvia, Gaysynsky y Cappella (2021) “la información errónea sobre la salud en las redes sociales requiere con urgencia una mayor acción de quienes trabajan en la investigación y la práctica de la salud pública” para entender mejor estos conceptos definen a la “información errónea de la salud” como información sobre la salud que es falsa o no cuenta con un consenso científico que lo valide. Al analizar el contenido es importante tener en cuenta varios factores como el reconocimiento de entidades, el sentimiento y el contexto al momento de la moderación.

### **Base conceptual**

En la industria del software el contenido es la parte más importante, es el motivo por la que las personas visitan y encuentran los sitios web, es el éxito de un aplicativo de software independientemente de su infraestructura es el contenido.

### ***Contenido web***

El contenido web puede definirse como la parte principal de un aplicativo, es el texto y los recursos de multimedia que los usuarios valoran, hace algunos años atrás una pequeña parte de sitios web contaban con una página de bienvenida, sin embargo, esto provocaba que muchos usuarios abandonen el sitio web, dado que busca información específica y se siente abrumado al observar una sobre carga de información. en la actualidad gran parte se genera de manera dinámica.

Los aplicativos webs más populares como las redes sociales proporcionan la facilidad de subir contenido a los mismos usuarios de esta manera interactúan entre sí, existen una gran diversidad de plataformas educativas, informativas, entretenimiento, sociales, foros entre otros, en que el usuario es quien genera y comparte contenido (Kimin, 2021).

### ***Análisis del contenido***

El análisis del contenido es interpretar los datos de manera cualitativa y cuantitativa de un aplicativo web con diversos fines, como la búsqueda de contenido homofóbico, ofensivo, reseñas positivas o negativas acerca de un producto o servicio, promueva el odio hacia diversos grupos étnico, religiosos u orientación sexual, terrorismo, tendencias del mercado, incluso nos permite determinar el comportamiento del usuario o de un grupo de usuarios, para la realización del análisis de contenido los usuarios aceptan que se realice en los términos y condiciones o políticas de uso de la aplicación (Guerrero, 2021).



El resultado del análisis de contenido permite a las organizaciones tomar decisiones en función de la situación, unos de los ejemplos más comunes es la censura de publicaciones que promueven el terrorismo, de esta manera los aplicativos webs como redes sociales no son el medio para darse a conocer estos grupos irregulares. el contenido generado por los usuarios debe analizarse por parte de las empresas protegiendo los derechos humanos de las personas, la libertad de expresión y la privacidad alineándose a las normas internacionales que establecen ciertos límites, además se debe considerar el principio de la internet abierta (UNESCO, 2021).

### ***Contenido ofensivo***

Según la real academia española (RAE, 2020) define ofender como “humillar o herir el amor propio o la dignidad de alguien, o ponerlo en evidencia con palabras o con hechos”, es común encontrar contenido con estas características y con el uso de palabras despectivas en sitios web populares como las redes sociales, según el instituto nacional de estadística y geografía menciona que el 40% de los usuarios de sitios web en general han sido ofendidos por medio de una publicación o mensaje escrito por otro usuario (INEGI, 2019).

Dada la libertad que los usuarios poseen para publicar contenido en foros, redes sociales y sitios web en general, existe un cierto abuso de la libertad de expresión, al escribir mensajes con texto despectivo y ofensivo. Uno de los tipos de contenido es el texto escrito, es el más usado dado que permite expresar ideas por medio de mensajes, comentarios o publicaciones sobre un determinado tema, sin embargo, el mensaje del contenido puede ofender a un grupo social, lo cual se debe evitar para crear un ambiente digno y armónico en la comunidad.

### ***Sentimientos***

La real academia española (RAE, 2020) define sentimiento como: “estado efectivo del ánimo”, existen tres sentimientos que predominan: positivo, neutral y negativo. con el primero el usuario expresa un mensaje de bienestar y sensación de agrado, en el segundo no expresa felicidad, pero tampoco existe una percepción de malestar y con el ultimo sentimiento existe una percepción de malestar o desagrado respecto algún tema o situación. a pesar de la existencia de muchos más sentimientos como la euforia, ira, miedo, tristeza, culpa entre otros. los tres ya mencionados (positivo, neutro, negativo) tienen mayor relevancia al momento de realizar un análisis de comentarios a gran escala.

### ***Entidades***

Las entidades corresponden a la información específica de personas, lugares, ubicaciones, expresiones numéricas como: fechas, tiempo y cantidades, personajes influyentes o temas en general que poseen un reconocimiento propio o único, un ejemplo común es el nombre de una ciudad (Guayaquil entidad de ciudad), otro tipo de entidad son la entidad nombrada se encuentran de manera explícita en texto no estructurado para los cuales existe un designador rígido (designa o refiere atributos propios) incluyendo sustantivos propios o términos para sustancias (Megias, s.f.).

El reconocimiento de entidades está conformado por dos partes, la primera es la obtención del atributo propio de la entidad como es el nombre y la segunda es la clasificación de la entidad según el tipo que sea, la empresa BBN technologies ha clasificado en 29 tipos, se debe considerar que la importancia del reconocimiento de las entidades está en su clasificación dependiendo el tipo.

### **Base legal**

La Constitución de la Republica del Ecuador (CONSTITUCION DE LA REPUBLICA DEL ECUADOR 2008) en el capítulo 6 los derechos de libertad en el artículo 66 en el literal 3 inciso a), reconoce y garantiza a las personas la integridad personal, que incluye: la física, psíquica, moral y sexual, en le literal 4 derecho a la igualdad formal, igualdad material y no discriminación. además, se fomenta y garantiza la libertad de expresión mediante los artículos 39 se garantiza los derechos de los jóvenes incluido la libertad de expresión. en la sección séptima comunicación social en el artículo 384 se asegura el ejercicio de los derechos de comunicación y libertad de expresión.

En el artículo 341 menciona el estado generará las condiciones para la protección integral de sus habitantes a lo largo de sus vidas, que aseguren los derechos y principios reconocidos en la constitución, en particular la igualdad en la diversidad y la no discriminación, y priorizará su acción hacia aquellos grupos que requieran consideración especial por la persistencia de desigualdades.

En el año 2020 surgió el proyecto de ley llamado “PROYECTO DE LEY ORGÁNICA LEY PARA PREVENIR LA VIOLENCIA, EL ACOSO DIGITAL Y LA VIOLACIÓN A LA INTIMIDAD” en el que se pide la modificación de artículos del código orgánico penal con referencia al acoso digital que existen en diversas plataformas web (Anibal, 2020).

### **Antecedentes Contextuales**

Jovenesweb es una plataforma de educación con presencia online (sitio web y aplicativo móvil) que comenzó a finales del 2016 y permite a los estudiantes de tercero de bachillerato o a las personas que ya se graduaron del colegio prepararse para acceder a la educación superior.

Esta plataforma cuenta con más de 1 millón de usuarios registrados y con más de 1000 usuarios activos por día, sus estudiantes acceden desde distintos países, aunque principalmente se destacan en Ecuador. Aproximadamente el 69% de los estudiantes que se preparan son mujeres mientras que el 31% restante son hombres y más del 50% de los estudiantes tienen un rango de edad que va entre los 18 a 34 años.

Para la justificación científica del problema se ha elaborado una entrevista al encargado de Jovenesweb, de la cual se ha obtenido los siguientes resultados.

Pregunta 1: ¿Con qué frecuencia analizan el contenido generado por los usuarios en Jovenesweb?

Depende mucho de la actividad de los usuarios en nuestra plataforma educativa, pero por lo general se realiza este análisis 2 o 3 veces por semana.

Pregunta 2: ¿Tienen personal encargado de la revisión del contenido?

Sí, por el momento tenemos a 2 personas encargadas de estar revisando el contenido en Jovenesweb.

Pregunta 3: ¿Usan alguna herramienta para la moderación de contenido?

No, por el momento hacemos todo el proceso de forma manual, es decir, se analiza cada comentario, respuesta o aportes generados por los estudiantes uno por uno y en caso de no ser apropiado se elimina el contenido de nuestra plataforma.

Pregunta 4: ¿Cuánto tiempo dedican a analizar el contenido en Jovenesweb?

Por lo general, entre 4 a 5 horas dependiendo de la cantidad de contenido generado por los usuarios.

Pregunta 5: ¿En algún momento han tenido problemas analizando el contenido en Jovenesweb?

Sí, cuando hay mucha actividad de los usuarios por lo general no se alcanza a revisar todo el contenido por lo que se suele acumular y esto a su vez provoca que haya contenido sin revisar en la plataforma.

Pregunta 6: ¿Qué dificultades han presentado al momento de analizar el contenido en Jovenesweb?

En ocasiones por fallas humanas algunos comentarios o aportes de los usuarios que deben ser eliminados son aprobados lo cual genera problemas ya que se requiere realizar una segunda revisión, lo cual conlleva más tiempo y gastos económicos.

Pregunta 7: ¿Cómo cree usted que se podría mejorar el análisis de contenido en Jovenesweb?

Con el avance de la tecnología, se podría usar algún tipo de servicio que nos permita automatizar esta revisión de contenido. Hemos buscado algunas alternativas, sin embargo, la mayoría se enfocan en la revisión de contenido en inglés y en Jovenesweb la mayoría del contenido, por no decir todo, es en español.

Pregunta 8: Actualmente, ¿generan algún tipo de reportes sobre el análisis de contenido en Jovenesweb?

No, como la revisión de contenido la realizamos de manera manual el generar reportes nos tomaría más tiempo.

Pregunta 9: ¿Apoyaría la creación de una aplicación web que permita automatizar el análisis de contenido usando servicios de inteligencia artificial en Jovenesweb?

Sí, creo que con los avances que ha tenido la tecnología se podría realizar este proceso de manera automática y que además nos sería de mucha ayuda ya que nos permitiría analizar el contenido de una manera más rápida. Como mencioné anteriormente hemos buscado este tipo de herramientas sin embargo hasta el momento no hemos encontrado una buena alternativa.

Pregunta 10: ¿Estaría dispuesto a usar esta aplicación web para analizar el contenido en Jovenesweb?

Si esta aplicación web nos permite reducir el tiempo de revisión del contenido, la usaríamos con todo gusto.

En base a las preguntas realizadas en la entrevista con el encargado de Jovenesweb podemos evidenciar que actualmente realizan el análisis del contenido de manera manual, es decir, revisan uno por uno los contenidos generados por los estudiantes, lo cual a su vez conlleva mucho tiempo de revisión y además este método puede presentar problemas ya que está expuesto a fallas humanas lo cual implicaría que deban revisar nuevamente el contenido para asegurarse de que la revisión es correcta.

También en la entrevista se pudo evidenciar que han tratado de buscar herramientas que les permita automatizar el análisis de contenido pero que no han podido encontrar una buena alternativa ya que la mayoría de las soluciones ofrecen la revisión de contenido, pero en inglés, algo que no les sería de ayuda ya que la mayoría del contenido es en español.

Finalmente, con la entrevista se pudo evidenciar el interés que se tiene en Jovenesweb para el desarrollo de una plataforma web que permita automatizar el proceso de análisis de contenido ya que así podrán reducir los tiempos de revisión.

## Capítulo III

### Desarrollo de la propuesta

#### **Etapa 1: Inicio del proyecto**

En el desarrollo de la etapa 1, se da inicio al proyecto definiendo la visión, el equipo scrum, el producto backlog, la arquitectura, de esta manera se establece las condiciones necesarias para el desarrollo del proyecto bajo un marco de trabajo ágil, además nos permitirá continuar con una forma de trabajo predictiva (certmind, 2020).

#### ***Visión del proyecto***

Para la plataforma educativa Jovenesweb, que maneja grandes cantidades de contenido generado por sus usuarios, Content Moderator es un aplicativo web que mediante el consumo de Servicios Cognitivos de Inteligencia Artificial analiza y clasifica dicho contenido en: sentimientos (positivo, negativo neutral), entidades (información específica de lugares, ubicaciones, expresiones numéricas, personajes influyentes ...) y si es ofensivo; mostrando los resultados del análisis mediante tablas interactivas y gráficos estadísticos que permiten tener un repositorio de información fiable para mejorar la toma de decisiones.

#### ***Equipo Scrum***

Durante esta etapa se eligen los miembros que formaran parte del equipo, las partes interesadas, quienes participaran y formaran parte del desarrollo del proyecto, cada uno tiene una función diferente entre los cuales predominaran los principios de Scrum para el éxito del proyecto y la mejora continua por parte del equipo.



## Core

Product Owner: Fernando Elvis Betancourt Ludeña

Scrum Master: Edgar Santiago Panata Castillo

Development Team:

- Elvis Fernando Betancourt Ludeña
- Edgar Santiago Panata Castillo

## Non-Core

Stakeholders: Jovenesweb

Scrum Guidance Body: Ing. Diego Ismael Velasco Sanchez

Estructura del equipo Scrum SBOK Guide published by SCRUMstudy (Shubhashree, 2018)

### ***Definición de roles del Product Backlog***

Cliente. - Administrador de Jovenesweb

### ***Product Backlog***

Parte de la primera etapa se genera el Product Backlog, el cual es una lista de requerimientos que el equipo de desarrollo debe desarrollar, los requerimientos son licitados mediante el uso de las historias de usuario, el objetivo de esta parte de la etapa es tener claro los requerimientos para cumplir con la visión del proyecto, se debe considerar las dependencias que puede existir en las diferentes historias de usuario.

## Tabla 2

### *Historia de usuario 1*

HU01	
Nombre	Página principal
<b>Como</b> cliente <b>quiero</b> que en la página principal se visualicen los servicios <b>para</b> que los usuarios puedan conocer los beneficios	

**Tabla 3***Historia de usuario 2*

<b>HU02</b>	
<b>Nombre</b>	Diseño de la base de datos
<b>Como</b> cliente <b>quiero</b> almacenar la información <b>para</b> mantener organizada la información	

**Tabla 4***Historia de usuario 3*

<b>HU03</b>	
<b>Nombre</b>	Creación de usuarios
<b>Como</b> cliente <b>quiero</b> crear una cuenta <b>para</b> acceder a los servicios	

**Tabla 5***Historia de usuario 4*

<b>HU04</b>	
<b>Nombre</b>	Inicio de sesión
<b>Como</b> cliente <b>quiero</b> acceder a mi cuenta con mi usuario y contraseña <b>para</b> administrar el contenido analizado	

**Tabla 6***Historia de usuario 5*

<b>HU05</b>	
<b>Nombre</b>	Categorizar contenido ofensivo
<b>Como</b> cliente <b>quiero</b> categorizar el contenido ofensivo <b>para</b> visualizar los <u>datos</u> de una mejor manera.	

**Tabla 7***Historia de usuario 6*

<b>HU06</b>	
<b>Nombre</b>	Agrupar la información por tipo de entidades
<b>HU06</b>	
<b>Como</b> cliente <b>quiero</b> categorizar el contenido por tipos de entidades <b>para</b> visualizar los datos de una mejor manera.	

**Tabla 8***Historia de usuario 7*

<b>HU07</b>	
<b>Nombre</b>	Creación de API
<b>Como</b> cliente <b>quiero</b> que exista un grupo de servicios como API <b>para</b> poder integrarlos en mi sistema.	

**Tabla 9***Historia de usuario 8*

<b>HU08</b>	
<b>Nombre</b>	Documentación de la API
<b>Como</b> cliente <b>quiero</b> que exista una documentación sobre el uso de la API <b>para</b> que los clientes conozcan la funcionalidad.	

**Tabla 10***Historia de usuario 9*

<b>HU09</b>	
<b>Nombre</b>	Reporte PDF de las entidades y contenido ofensivo
<b>Como</b> cliente <b>quiero</b> que se generen reportes en formato PDF de cada tipo de entidad y del contenido ofensivo <b>para</b> imprimirlos y llevar un registro del contenido analizado.	

**Tabla 11***Historia de usuario 10*


---

<b>HU10</b>	
<b>Nombre</b>	Reporte Excel de las entidades y contenido ofensivo
<b>Como</b> cliente <b>quiero</b> que se generen reportes en Excel de cada tipo de entidad y del contenido ofensivo <b>para</b> llevar un registro del contenido analizado.	

---

**Tabla 12***Historia de usuario 11*


---

<b>HU11</b>	
<b>Nombre</b>	Visualizar gráficas sentimientos
<b>Como</b> cliente <b>quiero</b> visualizar gráficos estadísticos del análisis de sentimientos <b>para</b> evaluar los sentimientos de los usuarios de Jovenesweb con respecto al contenido.	

---

**Tabla 13***Historia de usuario 12*


---

<b>HU12</b>	
<b>Nombre</b>	Visualizar gráficas contenido ofensivo
<b>Como</b> cliente <b>quiero</b> visualizar gráficos estadísticos del contenido ofensivo <b>para</b> determinar la tendencia de los usuarios de Jovenesweb.	

---

**Tabla 14***Historia de usuario 13*


---

<b>HU13</b>	
<b>Nombre</b>	Visualizar gráficas de categorías
<b>Como</b> cliente <b>quiero</b> visualizar las categorías más repetidas <b>para</b> conocer los temas con mayor interés en los usuarios de Jovenesweb.	

---

**Tabla 15***Historia de usuario 14*

HU14	
Nombre	Filtrar el contenido analizado
<b>Como</b> cliente <b>quiero</b> gestionar los resultados obtenidos por filtros y/o agrupaciones <b>para</b> tener una mejor comprensión de los resultados.	

***Priorización del Backlog***

La priorización se realizará con la técnica valor de negocio vs complejidad se debe considerar factores externos que influyen en el desarrollo del proyecto tales como: el tiempo, valor funcional, tecnología que se va a desarrollar, riesgos y esfuerzo, opinión por parte de los expertos en la lógica del negocio, al momento de dar un valor por parte del equipo de desarrollo. Con el uso de esta técnica se da un valor de negocio entre 1 al 10 por el producto Owner y las unidades de estimación de complejidad por parte del equipo de desarrollo, la priorización del producto backlog se puede observar en la tabla 16 (Scrum, 2020).

**Tabla 16***Product backlog priorizado*

Código	Valor de negocio	Puntos de historia	Valor de prioridad	Orden
HU01	7	1	7	1
HU02	6	4	1,5	2
HU03	6	4	1,5	3
HU04	8	7	1,14	5
HU05	8	9	0,88	8
HU06	7	6	1,16	4
HU07	4	5	0,8	12
HU08	5	6	0,83	11
HU09	5	7	0,71	13
HU10	7	8	0,875	10

<b>Código</b>	<b>Valor de negocio</b>	<b>Puntos de historia</b>	<b>Valor de prioridad</b>	<b>Orden</b>
HU11	7	8	0,875	9
HU12	7	7	1	6
HU13	6	6	1	7
HU14	4	6	0.66	14

### ***Arquitectura del proyecto***

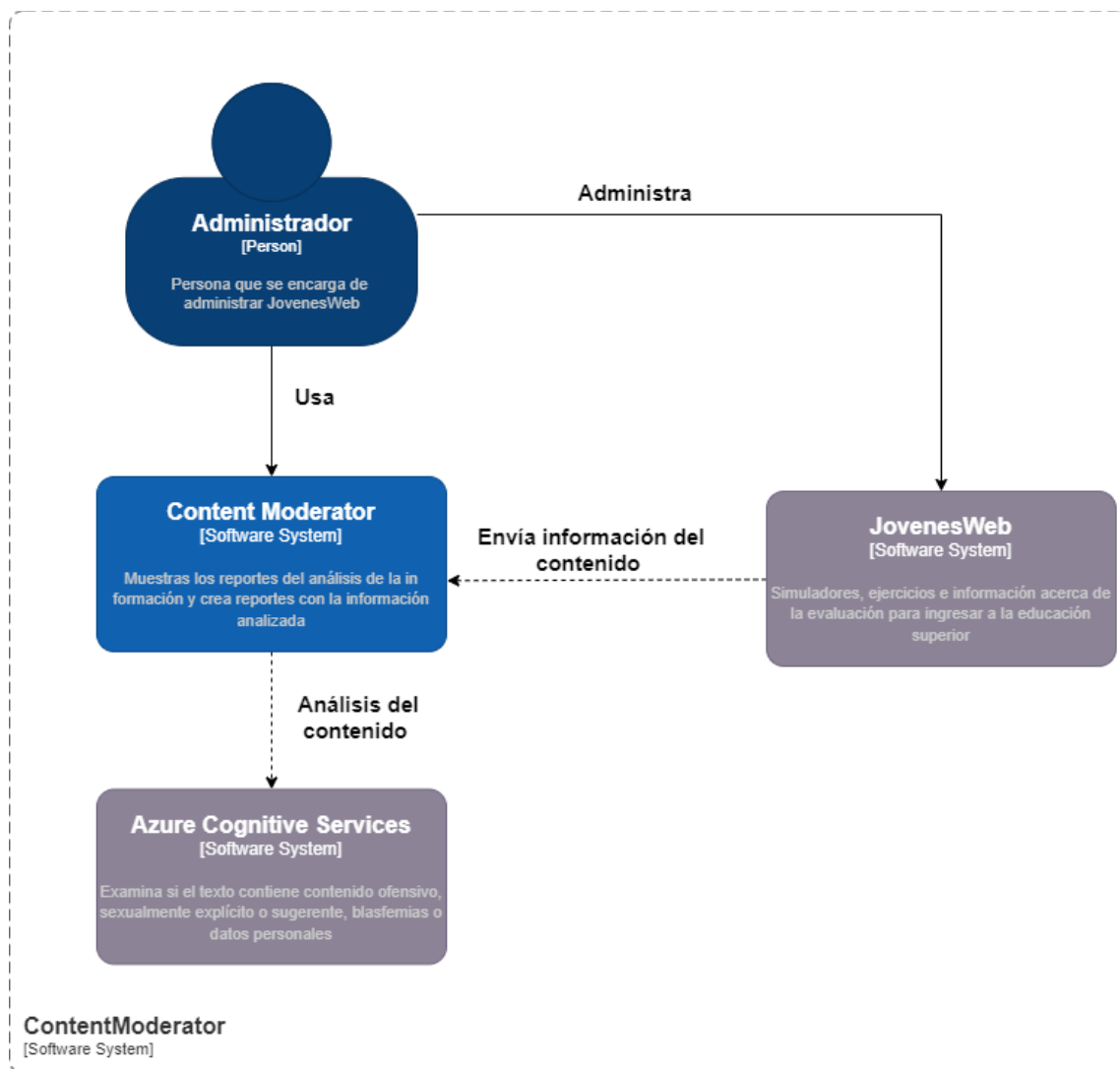
Mediante el diseño de la arquitectura del proyecto se refina el diseño técnico del producto, se desarrolla en la etapa inicial del proyecto para tener un enfoque global acerca de las herramientas y la comunicación que van existir entre ellas, la mayoría de arquitecturas requieren el conocimiento de una gran cantidad de símbolos, líneas, cajas y un conocimiento de UML, gran parte de los equipos prefieren diagramar ad hoc, pero esto podría traer inconvenientes al ser diagramas ambiguos, entonces se crea el modelo C4 para el diseño de la arquitectura, el cual se realiza en diferentes niveles y abstracciones (Vivanco, 2019).

#### **Nivel 1**

Diagrama de contexto o también conocido como nivel 1, muestra como el aplicativo web y como interactúa con las personas que lo utilizan y con los otros sistemas software, el objetivo de la abstracción es lograr que todos entiendan el diagrama incluyendo el cliente (muchos de los casos tienen un conocimiento técnico limitado)

**Figura 3**

Nivel 1 de la arquitectura del proyecto

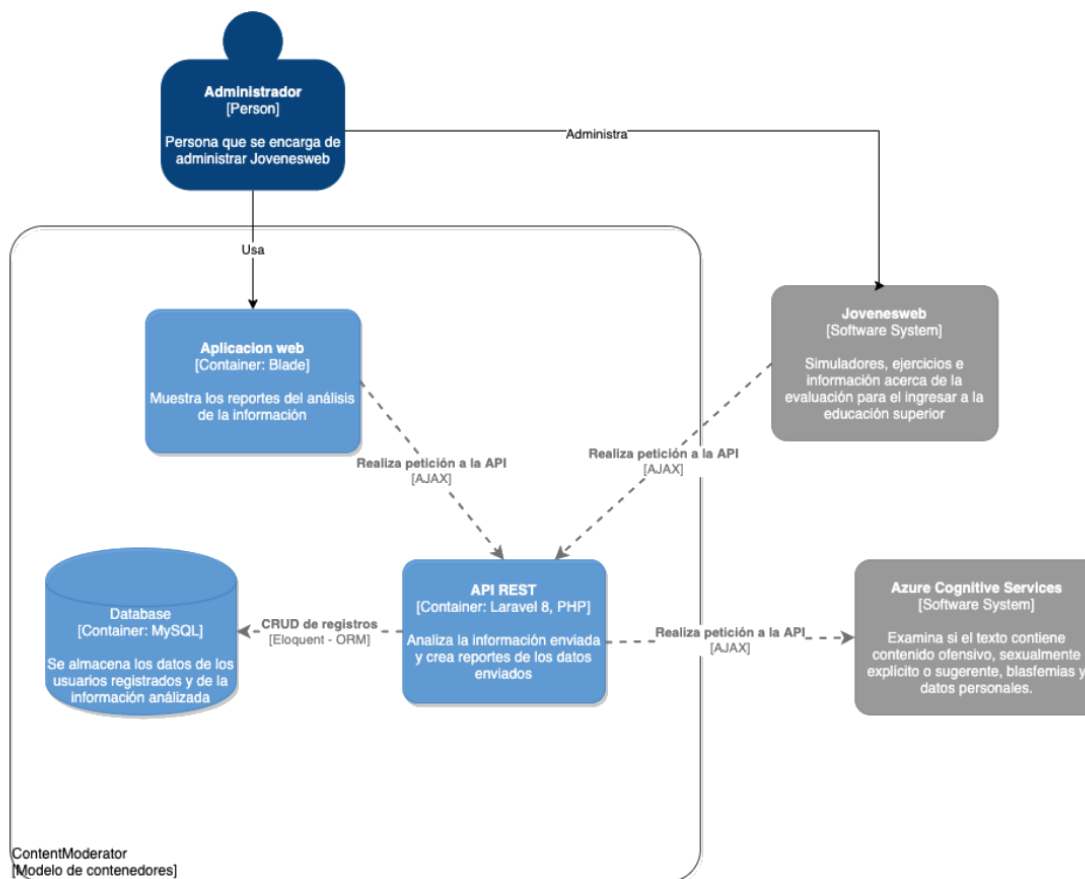


## Nivel 2

Diagrama de contenedor o también conocido como nivel 2, hace un acercamiento al aplicativo web mediante contenedores haciendo una breve descripción de las tecnologías que se están utilizando (lenguaje de programación, ORM, frameworks, base de datos), además de los diferentes protocolos de comunicación entre el aplicativo web desarrollado y los otros sistemas que interactúan

**Figura 4**

*Nivel 1 de la arquitectura del proyecto*



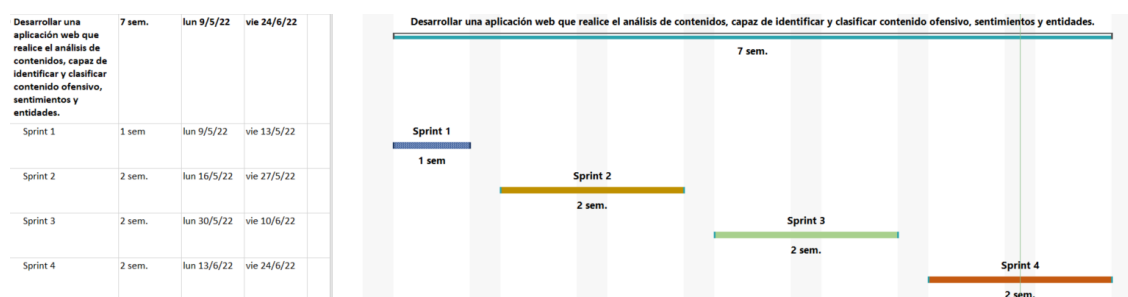
### **Cronograma de entregas**

El product owner de manera conjunta con el equipo de desarrollo definen el cronograma de entregas, el mismo está detallado la duración estimada de cada sprint en función a la lista del priorizada del product backlog, en la imagen # se detalla a un alto nivel los diferentes sprint



**Figura 5**

*Cronograma de entrega de los sprint a realizar*



*Nota.* Esta figura describe el cronograma de entrega de cada uno de los sprints

## **Etapa 2: Planificación del Sprint**

En este evento se socializa con el equipo de desarrollo las diferentes historias de usuario que se van a desarrollar, además de definir la meta del sprint y el tiempo en que se va a desarrollar, esta planificación está limitada por un tiempo específico de una hora por semana de iteración de cada sprint (Laoyan, 2021).

### ***Primer Sprint Planning***

**Tabla 17**

*Planning del primer sprint*

<b>Código</b>	<b>Descripción</b>	<b>Valor de estimación (días)</b>
HU01	Página principal	1
HU02	Diseño de la base de datos	2
HU03	Creación de usuarios	2
HU04	Inicio de sesión	2

*Nota.* Esta tabla describe la planificación del primer Sprint con la valoración de estimación

En la Tabla 17, se describe la planificación del primer sprint, donde se realiza la página principal de Content Moderator, aquí se expondrá una breve información sobre los servicios que ofrece, además se realizará el sistema de autenticación el cual consta de la creación de nuevos usuarios y del acceso a través de un usuario y contraseña.

### **Segundo Sprint Planning**

**Tabla 18**

*Planning del segundo sprint*

<b>Código</b>	<b>Descripción</b>	<b>Valor de estimación (días)</b>
HU05	Categorizar contenido ofensivo	4
HU06	Agrupar la información por tipo de entidades	4
HU07	Creación de API	3
HU08	Documentación de la API	3

*Nota.* Esta tabla describe la planificación del segundo Sprint con la valoración de estimación

En la Tabla 18 se evidencia la planificación del Sprint 2, donde se procederá a realizar la categorización del contenido ofensivo y agrupación de las entidades, lo que implica integrar los servicios cognitivos de inteligencia artificial de Azure con Content Moderator y a través de la creación de la API se almacenará la información analizada en la base de datos para posteriormente hacer uso de esta información, finalmente se realizará la documentación de cada URL de la API que contará con descripción, los parámetros para la petición y un ejemplo de cómo realizar la petición.

### **Tercer Sprint Planning**

**Tabla 19**

*Planning del tercer sprint*

<b>Código</b>	<b>Descripción</b>	<b>Valor de estimación (días)</b>
HU09	Reporte PDF de las entidades y contenido ofensivo	5
HU010	Reporte Excel de las entidades y contenido ofensivo	5
HU11	Visualizar gráficas sentimientos	4

*Nota.* Esta tabla describe la planificación del tercer Sprint con la valoración de estimación

En la Tabla 19, se visualiza la planificación del tercer Sprint, el cual consiste en la elaboración de reportes de la información analizada, dichos reportes serán en formato PDF y Excel y se podrán descargar por cada tipo de entidad o por contenido ofensivo, además se realizará el primer gráfico estadístico para visualizar los sentimientos analizados.

### ***Cuarto Sprint Planning***

**Tabla 20**

*Planning del cuarto sprint*

<b>Código</b>	<b>Descripción</b>	<b>Valor de estimación (días)</b>
HU12	Visualizar gráficas contenido ofensivo	4
HU13	Visualizar gráficas de categorías	4
HU14	Filtrar el contenido analizado	6

*Nota.* Esta tabla describe la planificación del cuarto Sprint con la valoración de estimación

Finalmente, en la Tabla 20 se realiza la planificación del cuarto Sprint, donde se elabora los gráficos estadísticos de la información del contenido ofensivo y las categorías, además se realizará la filtración del contenido el cual consiste en ordenar el contenido de forma ascendente o descendente o a su vez que se permita buscar el contenido para facilitar la búsqueda de la información.

### **Etapa 3: Desarrollo del Sprint**

El Desarrollo del sprint se lo realiza de manera interactiva teniendo en cuenta que la característica principal del sprint es el trabajo de manera paralela entre el desarrollo de los entregables y el scrum diario, además el uso de artefactos como el burndown chart del sprint nos permite medir la productividad del equipo entre el trabajo pendiente y días del sprint

**Sprint 1****Tabla 21***Criterios de aceptación de la historia de usuario 1*

<b>HU01</b>		
<b>Nombre</b>	<b>Página principal</b>	
<b>Como</b> cliente <b>quiero</b> que en la página principal se visualicen los servicios <b>para</b> que los usuarios puedan conocer los beneficios		
<b>Criterios de aceptación</b>		
<b>1</b>	Información principal	<p><b>Dado</b> que el cliente quiere saber para qué sirve Content Moderator</p> <p><b>Cuando</b> ingresa al sitio web</p> <p><b>Entonces</b> visualizará una descripción sobre la funcionalidad y sus características</p>
<b>2</b>	Iniciar sesión	<p><b>Dado</b> que el cliente quiere iniciar sesión en el sistema</p> <p><b>Cuando</b> ingresa a la página principal</p> <p><b>Entonces</b> verá un botón que lo redireccione a la página de inicio de sesión</p>
<b>3</b>	Crear cuenta	<p><b>Dado</b> que el cliente quiere crear una cuenta en el sistema</p> <p><b>Cuando</b> ingresa a la página principal</p> <p><b>Entonces</b> verá un botón que lo redireccione a la página de registro</p>
<b>4</b>	Acceder al dashboard	<p><b>Dado</b> que el cliente ya tiene la sesión abierta</p> <p><b>Cuando</b> está en la página principal</p> <p><b>Entonces</b> verá un botón que lo redireccione al dashboard</p>
<b>5</b>	Diseño responsive	<p><b>Dado</b> que el cliente quiere acceder desde su celular</p> <p><b>Cuando</b> está en la página principal</p> <p><b>Entonces</b> vera un diseño responsive adaptado al tamaño del dispositivo</p>

Tabla 22

*Criterios de aceptación de la historia de usuario 2*

HU02	
Nombre	Diseño de la base de datos
<b>Como</b> cliente <b>quiero</b> almacenar la información <b>para</b> mantener organizada la información	
Criterios de aceptación	
1	<p>Diseño relacional de la base de datos</p> <p><b>Dado</b> que el cliente menciona la relación de la información</p> <p><b>Cuando</b> almacena la información</p> <p><b>Entonces</b> se realiza un diseño relacional acorde a la información.</p>
2	<p>Implementación de la base de datos</p> <p><b>Dado</b> que el cliente quiere almacenar la información</p> <p><b>Cuando</b> se realice el análisis de contenido</p> <p><b>Entonces</b> se integra con las herramientas para el desarrollo</p>

Tabla 23

*Criterios de aceptación de la historia de usuario 3*

HU03	
Nombre	Creación de usuarios
<b>Como</b> cliente <b>quiero</b> crear una cuenta <b>para</b> acceder a los servicios	
Criterios de aceptación	
1	<p>Registro de nueva cuenta</p> <p><b>Dado</b> que el cliente quiere crear una cuenta</p> <p><b>Cuando</b> ingresa a la página de registro</p> <p><b>Entonces</b> se muestra un formulario con los campos de entrada: nombres, correo electrónico, contraseña y confirmación de contraseña.</p>

HU03		
2	Enlace página de iniciar sesión	<p><b>Dado</b> que el cliente ya tiene una cuenta</p> <p><b>Cuando</b> ingresa a la página de registro</p> <p><b>Entonces</b> verá un enlace que lo redirigirá a la página de inicio de sesión</p>
3	Validación del registro	<p><b>Dado</b> que el cliente ingresa la información incorrecta</p> <p><b>Cuando</b> está en la página de registro</p> <p><b>Entonces</b> verá un mensaje de validación de los campos incorrectos</p>
4	Confirmación de contraseña	<p><b>Dado</b> que el cliente ingresa la contraseña</p> <p><b>Cuando</b> está en la página de registro</p> <p><b>Entonces</b> deberá confirmar la contraseña en nuevo campo.</p>
5	Completar registro	<p><b>Dado</b> que el cliente llena el formulario de registro</p> <p><b>Cuando</b> da clic en el botón de registro</p> <p><b>Entonces</b> se redirige a la página del dashboard con la sesión iniciada.</p>

**Tabla 24**

*Criterios de aceptación de la historia de usuario 4*

HU04		
Nombre		Inicio de sesión
<b>Como</b> cliente <b>quiero</b> acceder a mi cuenta con mi usuario y contraseña <b>para</b> administrar el contenido analizado		
Criterios de aceptación		
1	Iniciar sesión	<p><b>Dado</b> que el cliente quiere iniciar sesión</p> <p><b>Cuando</b> ingresa a la página de inicio de sesión</p> <p><b>Entonces</b> se muestra un formulario con los campos de entrada: correo electrónico y contraseña</p>
2	Enlace página de registro	<p><b>Dado</b> que el cliente no tiene una cuenta</p> <p><b>Cuando</b> ingresa a la página de registro</p> <p><b>Entonces</b> verá un enlace que lo redirigirá a la página de crear una nueva cuenta.</p>

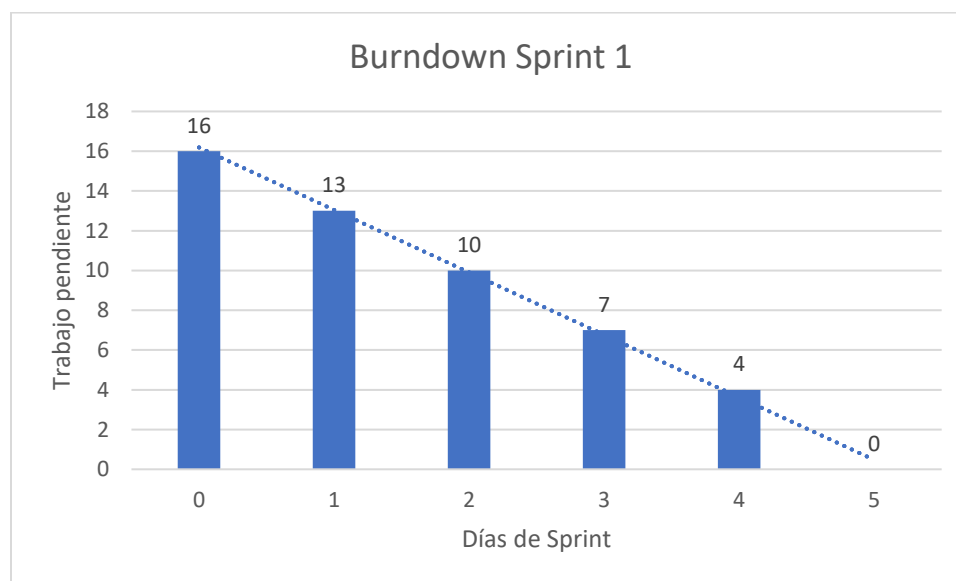
---

**HU04**


---

3	Validación del inicio de sesión	<p><b>Dado</b> que el cliente ingresa la información incorrecta</p> <p><b>Cuando</b> está en la página de inicio de sesión</p> <p><b>Entonces</b> verá un mensaje de validación de los campos incorrectos.</p>
4	Recordar sesión	<p><b>Dado</b> que el cliente quiere guardar su sesión por más tiempo</p> <p><b>Cuando</b> está ingresa su usuario y contraseña</p> <p><b>Entonces</b> deberá seleccionar el campo “Recordarme”.</p>
4	Confirmar inicio de sesión	<p><b>Dado</b> que el cliente ingresa su usuario y contraseña</p> <p><b>Cuando</b> da clic en el botón acceder</p> <p><b>Entonces</b> se redirige a la página del dashboard con su sesión iniciada.</p>

---

**Figura 6**
*Tablero de productividad del Sprint 1*


**Sprint 2****Tabla 25***Criterios de aceptación de la historia de usuario 5*

<b>HU05</b>	
<b>Nombre</b>	Contenido ofensivo
<b>Como</b> cliente <b>quiero</b> categorizar el contenido ofensivo <b>para</b> visualizar los <u>datos</u> de una mejor manera.	
<b>Criterios de aceptación</b>	
<b>1</b>	<p>Listado de contenido ofensivo</p> <p><b>Dado</b> que el cliente quiere visualizar el contenido analizado</p> <p><b>Cuando</b> ingresa a la sección de contenido ofensivo</p> <p><b>Entonces</b> se visualizará la información de cada texto donde se especifica si es ofensivo o no.</p>
<b>2</b>	<p>Paginación información</p> <p><b>Dado</b> que el cliente quiere visualizar el contenido analizado</p> <p><b>Cuando</b> ingresa a la sección de contenido ofensivo</p> <p><b>Entonces</b> si por defecto existen más de 10 registros se mostrará los botones de paginación.</p>
<b>3</b>	<p>Cantidad de registros por pagina</p> <p><b>Dado</b> que el cliente quiere visualizar el contenido analizado</p> <p><b>Cuando</b> ingresa a la sección de contenido ofensivo</p> <p><b>Entonces</b> podrá seleccionar la cantidad entre 10, 25, 50 o 100 registros por página.</p>
<b>4</b>	<p>Reporte PDF</p> <p><b>Dado</b> que el cliente quiere un reporte del contenido analizado</p> <p><b>Cuando</b> ingresa a la sección de contenido ofensivo</p> <p><b>Entonces</b> podrá descargar un reporte en PDF de la información.</p>
<b>5</b>	<p>Reporte Excel</p> <p><b>Dado</b> que el cliente quiere un reporte del contenido analizado</p> <p><b>Cuando</b> ingresa a la sección de contenido ofensivo</p>



<b>HU05</b>
<b>Entonces</b> podrá descargar un reporte en Excel de la información.

**Tabla 26**

*Criterios de aceptación de la historia de usuario 6*

<b>HU06</b>	
<b>Nombre</b>	Agrupar la información
<b>Como</b> cliente <b>quiero</b> categorizar el contenido por tipos de entidades <b>para</b> visualizar los <u>datos</u> de una mejor manera.	
<b>Criterios de aceptación</b>	
1	<p>Listado de categorías y subcategorías</p> <p><b>Dado</b> que el cliente quiere visualizar el contenido analizado</p> <p><b>Cuando</b> ingresa a la sección de categorías</p> <p><b>Entonces</b> se visualizará la información del contenido con su categoría y su subcategoría.</p>
2	<p>Paginación de categorías y subcategorías</p> <p><b>Dado</b> que el cliente quiere visualizar el contenido analizado</p> <p><b>Cuando</b> ingresa a la sección de categorías</p> <p><b>Entonces</b> se habilitarán los botones de paginación si existen más de 10 registros.</p>
3	<p>Cantidad de registros por página de categorías y subcategorías</p> <p><b>Dado</b> que el cliente quiere visualizar el contenido analizado</p> <p><b>Cuando</b> ingresa a la sección de categorías</p> <p><b>Entonces</b> podrá seleccionar la cantidad entre 10, 25, 50 o 100 registros por página.</p>
4	<p>Listado de categorías y subcategorías</p> <p><b>Dado</b> que el cliente quiere visualizar el contenido analizado</p> <p><b>Cuando</b> ingresa a la sección de información personal</p> <p><b>Entonces</b> se visualizará la información de cada entidad y la categoría.</p>

---

HU06		
5	Paginación del listado de información personal	<p><b>Dado</b> que el cliente quiere visualizar el contenido analizado</p> <p><b>Cuando</b> ingresa a la sección de información personal</p> <p><b>Entonces</b> se habilitarán los botones de paginación si existen más de 10 registros.</p>
6	Cantidad de registros por página de información personal	<p><b>Dado</b> que el cliente quiere visualizar el contenido analizado</p> <p><b>Cuando</b> ingresa a la sección de información personal</p> <p><b>Entonces</b> podrá seleccionar la cantidad entre 10, 25, 50 o 100 registros por página.</p>
7	Listado de enlaces	<p><b>Dado</b> que el cliente quiere visualizar el contenido analizado</p> <p><b>Cuando</b> ingresa a la sección de enlaces</p> <p><b>Entonces</b> se visualizará la información de cada entidad, la fuente y su URL</p>
8	Paginación del listado de enlaces	<p><b>Dado</b> que el cliente quiere visualizar el contenido analizado</p> <p><b>Cuando</b> ingresa a la sección de enlaces</p> <p><b>Entonces</b> se habilitarán los botones de paginación si existen más de 10 registros.</p>
9	Cantidad de registros por página de enlaces	<p><b>Dado</b> que el cliente quiere visualizar el contenido analizado</p> <p><b>Cuando</b> ingresa a la sección de enlaces</p> <p><b>Entonces</b> podrá seleccionar la cantidad entre 10, 25, 50 o 100 registros por página.</p>

---

Tabla 27

*Criterios de aceptación de la historia de usuario 7*

HU07		
Nombre		Creación de API
<b>Como</b> cliente <b>quiero</b> que exista un grupo de servicios como API <b>para</b> poder integrarlos en mi sistema.		
Criterios de aceptación		
1	API de categoría	<p><b>Dado</b> que el cliente desea integrar Jovenesweb con Content Moderator</p> <p><b>Cuando</b> necesita analizar las categorías de un contenido</p> <p><b>Entonces</b> dispondrá de una API para hacerlo</p>
2	API de información personal	<p><b>Dado</b> que el cliente desea integrar Jovenesweb con Content Moderator</p> <p><b>Cuando</b> necesita analizar la información personal de un contenido</p> <p><b>Entonces</b> dispondrá de una API para hacerlo</p>
3	Api de enlaces	<p><b>Dado</b> que el cliente desea integrar Jovenesweb con Content Moderator</p> <p><b>Cuando</b> necesita analizar los enlaces de un contenido</p> <p><b>Entonces</b> dispondrá de una API para hacerlo</p>
4	API de ofensivo	<p><b>Dado</b> que el cliente desea integrar Jovenesweb con Content Moderator</p> <p><b>Cuando</b> necesita analizar si un contenido es ofensivo</p> <p><b>Entonces</b> dispondrá de una API para hacerlo.</p>
5	API de sentimiento	<p><b>Dado</b> que el cliente desea integrar Jovenesweb con Content Moderator</p>

---

**HU07**


---

		<p><b>Cuando</b> necesita analizar los sentimientos de un contenido</p> <p><b>Entonces</b> dispondrá de una API para hacerlo</p> <p><b>Dado</b> que el cliente desea integrar Jovenesweb con Content Moderator</p>
<b>6</b>	API de categorías	<p><b>Cuando</b> necesita analizar las categorías de un conjunto de contenidos</p> <p><b>Entonces</b> dispondrá de una API para hacerlo</p> <p><b>Dado</b> que el cliente desea integrar Jovenesweb con Content Moderator</p>
<b>7</b>	API de informaciones personales	<p><b>Cuando</b> necesita analizar la información personal de un conjunto de contenidos</p> <p><b>Entonces</b> dispondrá de una API para hacerlo</p> <p><b>Dado</b> que el cliente desea integrar Jovenesweb con Content Moderator</p>
<b>8</b>	API de enlaces	<p><b>Cuando</b> necesita analizar los enlaces de un conjunto de contenidos</p> <p><b>Entonces</b> dispondrá de una API para hacerlo</p> <p><b>Dado</b> que el cliente desea integrar Jovenesweb con Content Moderator</p>
<b>9</b>	API de ofensivos	<p><b>Cuando</b> necesita analizar si un conjunto de contenidos es ofensivo</p> <p><b>Entonces</b> dispondrá de una API para hacerlo</p> <p><b>Dado</b> que el cliente desea integrar Jovenesweb con Content Moderator</p>
<b>10</b>	API de sentimientos	<p><b>Cuando</b> necesita analizar los sentimientos de un conjunto de contenidos</p> <p><b>Entonces</b> dispondrá de una API para hacerlo</p>

---

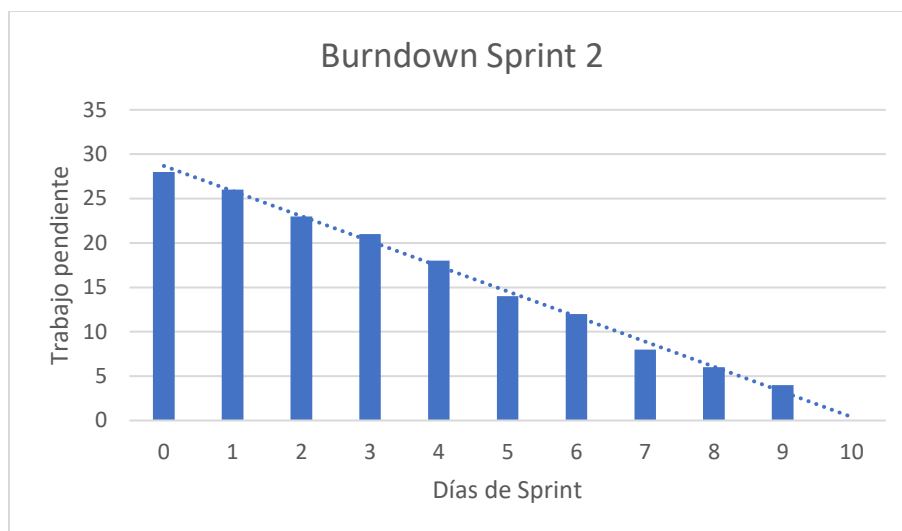
Tabla 28

*Criterios de aceptación de la historia usuario 8*

HU08	
Nombre	Documentación de la API
<p><b>Como</b> cliente <b>quiero</b> que exista una documentación sobre el uso de la API <b>para</b> que los clientes conozcan la funcionalidad.</p>	
Criterios de aceptación	
1	<p>Información de la API</p> <p><b>Dado</b> que el cliente desea conocer las distintas APIs disponibles y sus parámetros</p> <p><b>Cuando</b> quiere obtener detalles de la API</p> <p><b>Entonces</b> podrá ver las URLs y parámetros de cada API</p>
2	<p>Descripción de la API</p> <p><b>Dado</b> que el cliente desea conocer que hace cada API</p> <p><b>Cuando</b> necesita más información sobre el funcionamiento de la API</p> <p><b>Entonces</b> podrá ver una descripción de lo que hace cada API</p>
3	<p>Ejemplo de consumo de la API</p> <p><b>Dado</b> que el cliente desea conocer cómo integrar la API en su sistema</p> <p><b>Cuando</b> necesita va a hacer uso de la API</p> <p><b>Entonces</b> podrá ver un ejemplo de cómo realizar la petición a la API</p>
4	<p>Menú documentación</p> <p><b>Dado</b> que el cliente desea acceder a las distintas secciones</p> <p><b>Cuando</b> accede a la documentación</p> <p><b>Entonces</b> tendrá un menú en la parte izquierda para acceder a cada sección</p>

Figura 7

Tablero de productividad del Sprint 2



### Sprint 3

Tabla 29

Criterios de aceptación de la historia de usuario 9

HU09			
Nombre		Análisis de contenido reporte PDF	
<b>Como</b> cliente <b>quiero</b> que se generen reportes en formato PDF de cada tipo de entidad y del contenido ofensivo <b>para</b> imprimirlos y llevar un registro del contenido analizado.			
Criterios de aceptación			
1	Reporte de categorías	PDF de categorías	<p><b>Dado</b> que el cliente necesita obtener un reporte en PDF</p> <p><b>Cuando</b> este en el listado de categorías</p> <p><b>Entonces</b> podrá generar el reporte a través de un botón en la parte superior</p>
2	Reporte de enlaces	PDF de enlaces	<p><b>Dado</b> que el cliente necesita obtener un reporte en PDF</p> <p><b>Cuando</b> este en el listado de enlaces</p> <p><b>Entonces</b> podrá generar el reporte a través de un botón en la parte superior</p>

HU09	
3	<p>Reporte PDF de Información personal</p> <p><b>Dado</b> que el cliente necesita obtener un reporte en PDF</p> <p><b>Cuando</b> este en el listado de información personal</p> <p><b>Entonces</b> podrá generar el reporte a través de un botón en la parte superior</p>
4	<p>Reporte PDF de contenido ofensivo</p> <p><b>Dado</b> que el cliente necesita obtener un reporte en PDF</p> <p><b>Cuando</b> este en el listado de contenido ofensivo</p> <p><b>Entonces</b> podrá generar el reporte a través de un botón en la parte superior</p>

**Tabla 30**

*Criterios de aceptación de la historia de usuario 10*

HU10	
Nombre	Análisis de contenido reporte Excel
<p><b>Como</b> cliente <b>quiero</b> que se generen reportes en Excel de cada tipo de entidad y del contenido ofensivo <b>para</b> llevar un registro del contenido analizado.</p>	
Criterios de aceptación	
1	<p>Reporte Excel de categorías</p> <p><b>Dado</b> que el cliente necesita obtener un reporte en Excel</p> <p><b>Cuando</b> este en el listado de categorías</p> <p><b>Entonces</b> podrá generar el reporte a través de un botón en la parte superior</p>
2	<p>Reporte Excel de enlaces</p> <p><b>Dado</b> que el cliente necesita obtener un reporte en Excel</p> <p><b>Cuando</b> este en el listado de enlaces</p> <p><b>Entonces</b> podrá generar el reporte a través de un botón en la parte superior</p>
3	<p>Reporte Excel de Información personal</p> <p><b>Dado</b> que el cliente necesita obtener un reporte en Excel</p>

HU10	
4	Reporte Excel de contenido ofensivo
	<p><b>Cuando</b> este en el listado de información personal</p> <p><b>Entonces</b> podrá generar el reporte a través de un botón en la parte superior</p> <p><b>Dado</b> que el cliente necesita obtener un reporte en Excel</p> <p><b>Cuando</b> este en el listado de contenido ofensivo</p> <p><b>Entonces</b> podrá generar el reporte a través de un botón en la parte superior</p>

**Tabla 31**

*Criterios de aceptación de la historia de usuario 11*

HU11	
Nombre	Visualizar gráficas sentimientos
<b>Como</b> cliente <b>quiero</b> visualizar gráficos estadísticos del análisis de sentimientos <b>para</b> evaluar los sentimientos de los usuarios de Jovenesweb con respecto al contenido.	
Criterios de aceptación	
1	Gráfico cantidad de textos por sentimientos
	<p><b>Dado</b> que el cliente quiere visualizar la cantidad textos por sentimientos</p> <p><b>Cuando</b> ingresa a la página principal de cada grupo</p> <p><b>Entonces</b> visualizará un gráfico de barras de cantidad vs sentimientos</p>
2	Gráfico porcentual de sentimientos
	<p><b>Dado</b> que el cliente quiere visualizar el promedio porcentual de sentimientos</p> <p><b>Cuando</b> ingresa a la página principal de cada grupo</p> <p><b>Entonces</b> visualizará un gráfico circular con el promedio de los sentimientos predominantes analizados de cada texto</p>
3	Estandarización de colores de los sentimientos
	<p><b>Dado</b> que el cliente quiere diferenciar cada sentimiento en los gráficos</p> <p><b>Cuando</b> visualiza el gráfico circular o de barras de los sentimientos</p>



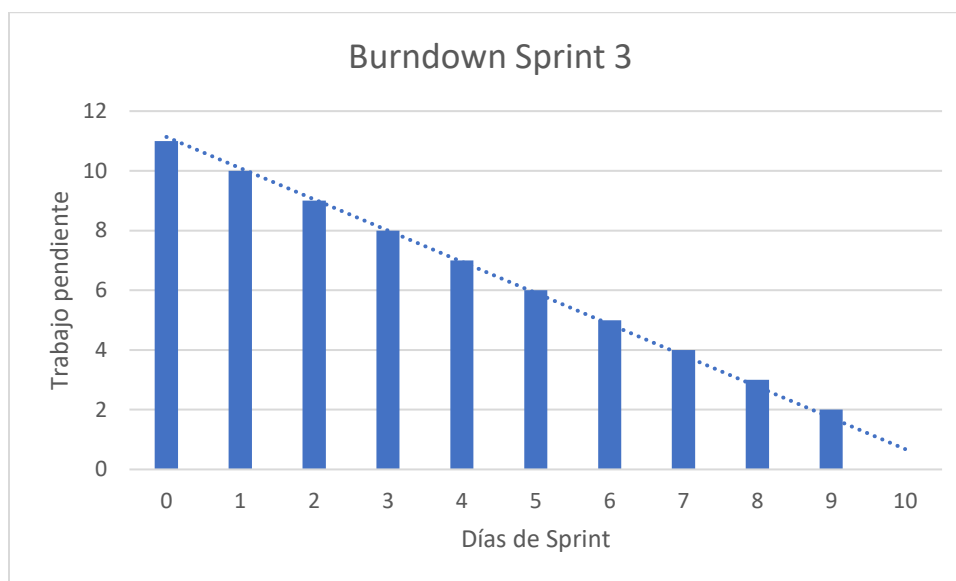
---

**HU11**


---

**Entonces** verá de color verde a los sentimientos positivos, de color amarillo al sentimiento neutral y de color rojo a los negativos

---

**Figura 8***Tablero de productividad del Sprint 3***Sprint 4****Tabla 32***Criterios de aceptación de la historia de usuario 12*


---

**HU12**


---

**Nombre**

Visualizar gráficas contenido ofensivo

**Como** cliente **quiero** visualizar gráficos estadísticos del contenido ofensivo **para** determinar la tendencia de los usuarios de Jovenesweb.

---

**Criterios de aceptación**

**Dado** que el cliente quiere conocer el porcentaje de

1      Porcentaje ofensivo      textos ofensivos

**Cuando** ingresa a la página principal de cada grupo

---

HU12	
	<p><b>Entonces</b> visualizará un gráfico circular del porcentaje de texto ofensivo y no ofensivo</p> <p><b>Dado</b> que el cliente quiere conocer las malas palabras más repetidas</p>
2	<p>Palabras repetidas</p> <p><b>Cuando</b> ingresa a la página principal de cada grupo</p> <p><b>Entonces</b> visualizará un gráfico de barras de las 10 palabras ofensivas más repetidas</p> <p><b>Dado</b> que el cliente quiere diferenciar el contenido ofensivo y no ofensivo en los gráficos</p>
3	<p>Estandarización de colores del contenido ofensivo</p> <p><b>Cuando</b> visualiza el gráfico circular del porcentaje de contenido ofensivo</p> <p><b>Entonces</b> verá de color rojo el contenido ofensivo y de color azul el contenido no ofensivo</p>

**Tabla 33**

Criterios de aceptación de la historia de usuario 13

HU13	
Nombre	Visualizar gráficas de categorías
<b>Como</b> cliente <b>quiero</b> visualizar las categorías más repetidas <b>para</b> conocer los temas con mayor interés en los usuarios de Jovenesweb.	
Criterios de aceptación	
1	<p>Cantidad de categorías en la gráfica</p> <p><b>Dado</b> que el cliente quiere visualizar las categorías más repetidas</p> <p><b>Cuando</b> ingresa a la página principal de cada grupo</p> <p><b>Entonces</b> visualizará solamente las 10 categorías más repetidas</p>
2	<p>Gráfica de barras de las categorías más repetidas</p> <p><b>Dado</b> que el cliente quiere visualizar las categorías más repetidas</p> <p><b>Cuando</b> ingresa a la página principal de cada grupo</p>

HU13	
	<p><b>Entonces</b> visualizará las categorías en una gráfica de barras del total de repeticiones vs el nombre de la entidad</p> <p><b>Dado</b> que el cliente quiere visualizar las categorías más repetidas</p>
3	<p>Estandarización de colores de las categorías</p> <p><b>Cuando</b> ingresa a la página principal de cada grupo</p> <p><b>Entonces</b> visualizará las barras del gráfico de color azul</p>

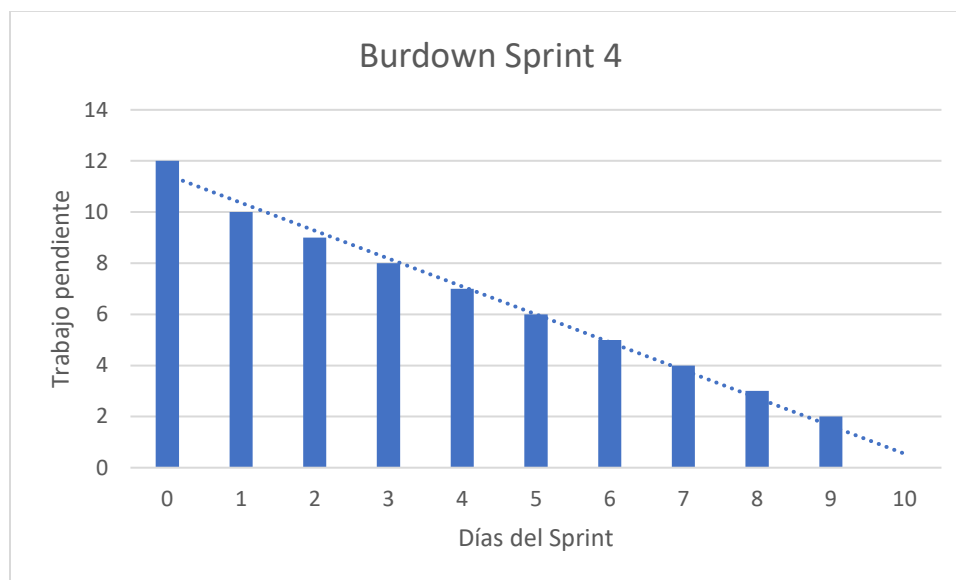
**Tabla 34**

*Criterios de aceptación de la historia de usuario 14*

HU14	
Nombre	Filtrar el contenido analizado
<p><b>Como</b> cliente <b>quiero</b> gestionar los resultados obtenidos por filtros y/o agrupaciones <b>para</b> tener una mejor comprensión de los resultados.</p>	
Criterios de aceptación	
1	<p>Filtrar y/o agrupar las categorías</p> <p><b>Dado</b> que el cliente quiere filtrar y/o agrupar</p> <p><b>Cuando</b> ingresa al listado de categorías</p> <p><b>Entonces</b> podrá visualizar una tabla que le permita ordenar de forma ascendente o de descendente cada columna y/o filtrar según el contenido</p>
2	<p>Filtrar y/o agrupar la información personal</p> <p><b>Dado</b> que el cliente quiere filtrar y/o agrupar</p> <p><b>Cuando</b> ingresa al listado de información personal</p> <p><b>Entonces</b> podrá visualizar una tabla que le permita ordenar de forma ascendente o de descendente cada columna y/o filtrar según el contenido</p>
3	<p>Filtrar y/o agrupar los enlaces</p> <p><b>Dado</b> que el cliente quiere filtrar y/o agrupar</p> <p><b>Cuando</b> ingresa al listado de enlaces</p> <p><b>Entonces</b> podrá visualizar una tabla que le permita ordenar de forma ascendente o de descendente cada columna y/o filtrar según el contenido</p>

## Figura 9

*Tablero de productividad del Sprint 4*



### **Etapa 4: Revisión del Sprint**

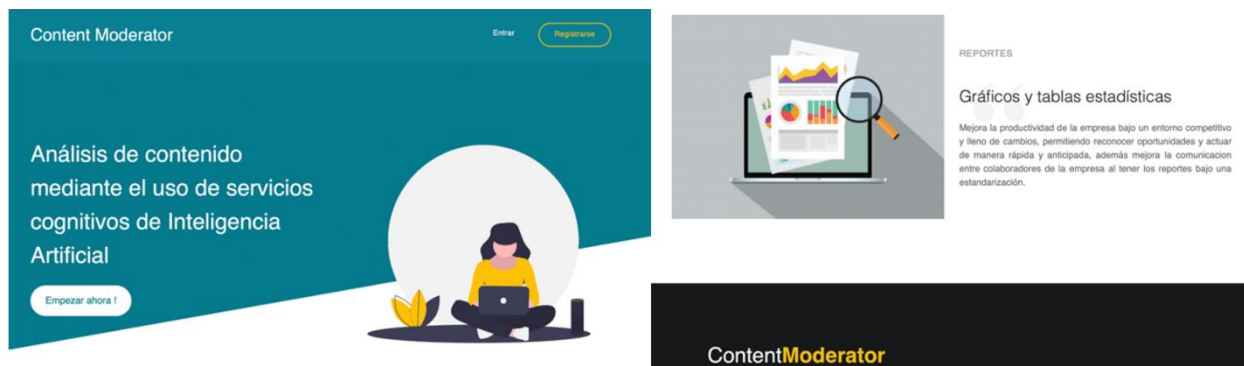
La revisión del sprint se realiza de manera interactiva al final de cada sprint desarrollado, durante este evento participa: producto Owner, Scrum Master, equipo de desarrollo y el cliente, en donde se revisa el entregable desarrollado y si se logró cumplir con la meta del sprint, el tiempo de duración es una 1 hora por semana de desarrollo, la participación por parte del cliente es activa y se hace uso de las historias de usuario para la revisión (Garcia M. , 2020).

#### ***Entregable Sprint 1***

El Sprint 1 con duración de 1 semana se realizó del 09/05/2022 hasta el 13/05/2022, en el cual se desarrollaron distintas tareas para dar cumplimiento a las historias de usuario HU01, HU02, HU03 y HU04, de las cuales a continuación se mostrarán los avances realizados:

## Figura 10

### *Página principal de Content Moderator*

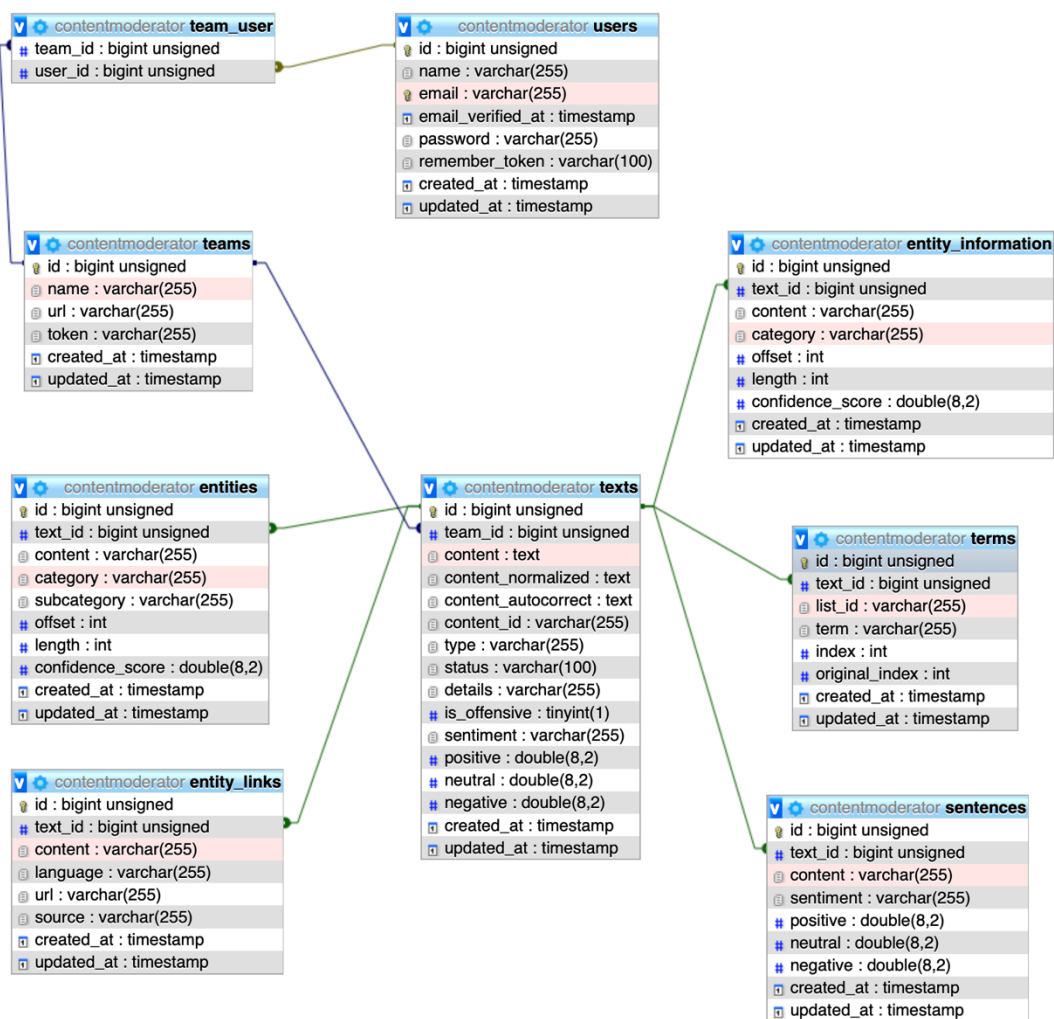


*Nota. En el gráfico representa el desarrollo de la página principal*

En la Figura 10, se puede evidenciar el desarrollo de la página principal de Content Moderator donde se muestra una descripción para que el cliente pueda conocer sobre su funcionalidad y características, además en la parte superior derecha se incluyen dos enlaces para que el cliente pueda crear una nueva cuenta o si ya tiene una, acceder con su usuario y contraseña. Además, se realizó el diseño para que sea responsive y se pueda visualizar correctamente en celulares.

## Figura 11

### *Diseño del modelo de la base de datos*



*Nota.* El gráfico representa el modelo relacional de Content Moderator

Para el desarrollo de Content Moderator se usó la base de datos relacional de MySQL en la cual a cada tabla se la nombró en inglés y en plural, esto en base a los estándares recomendados por el Framework Laravel. Además, se puede identificar a la tabla “Texts” como la más relevante ya que es la tabla que más se relaciona con las demás, tal y como se muestra en la Figura 2. Finalmente, se puede identificar dos tipos de relaciones para el diseño, la relación de Uno a Muchos (One-To-Many) que se realiza de la tabla “Texts” con las demás y la relación Muchos a Muchos (Many-To-Many) entre las tablas “Teams” y “Users”.

## Figura 12

*Página de creación de una cuenta*

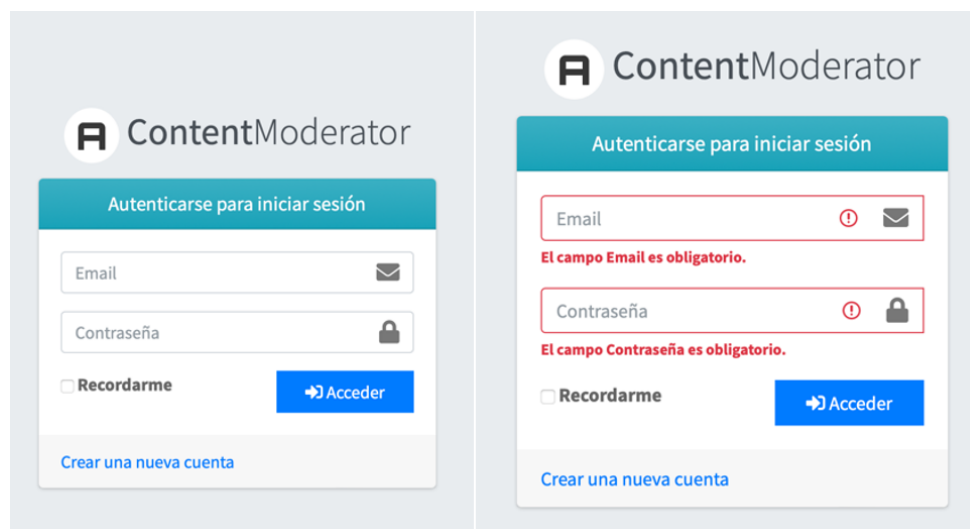
The image displays two versions of the 'ContentModerator' registration page. Both pages have a teal header with the 'ContentModerator' logo and the title 'Crear una nueva cuenta'. The left version shows a clean form with four input fields: 'Nombre completo' (with a person icon), 'Email' (with an envelope icon), 'Contraseña' (with a lock icon), and 'Confirmar la contraseña' (with a lock icon). Below the fields is a blue 'Registrarse' button and a link 'Ya tengo una cuenta'. The right version shows the same form but with red error messages: 'El campo Nombres completos es obligatorio.' under the first field, 'El campo Email es obligatorio.' under the second, and 'El campo Contraseña es obligatorio.' under the third. The 'Registrarse' button and the 'Ya tengo una cuenta' link are also present in this version.

*Nota.* El gráfico representa el formulario del registro principal

En la Figura 2, se puede evidenciar el desarrollo del formulario de registro en el cual están disponibles los campos Nombre completo, Email, Contraseña y Confirmar contraseña con su respectiva validación, además de un botón para registrar una nueva cuenta en base a los datos ingresados. También se agregó en la parte inferior un enlace para que el cliente pueda acceder al inicio de sesión en caso de que ya tenga una cuenta creada.

## Figura 13

### Página de inicio de sesión



The figure displays two versions of the ContentModerator login page. The left version is the standard login form, featuring the ContentModerator logo, a teal header with the text 'Autenticarse para iniciar sesión', and input fields for 'Email' and 'Contraseña'. Below these fields are a 'Recordarme' checkbox and a blue 'Acceder' button. A link for 'Crear una nueva cuenta' is located at the bottom. The right version shows the form after a validation error, with red boxes around the input fields and red text messages: 'El campo Email es obligatorio.' and 'El campo Contraseña es obligatorio.'.

*Nota.* El gráfico representa el formulario de inicio de sesión

En la Figura 13, se muestra el desarrollo realizado para el formulario de inicio de sesión, el cual contiene los campos Email y Contraseña y que cuentan cada uno con su respectiva validación, además se agregó la opción para que el usuario pueda escoger si desea tener la sesión abierta durante un mayor tiempo y finalmente se añadió un enlace para que el cliente se pueda dirigir a la página de registro en el caso de no estar registrado.

### **Entregable Sprint 2**


El Sprint 2 con duración de 2 semanas se realizó del 16/05/2022 hasta el 20/06/2022, en el cual se desarrollaron distintas tareas para dar cumplimiento a las historias de usuario HU08, HU09, y HU10, de las cuales a continuación se mostrarán los avances realizados:



## Figura 14

### Categorización de los registros ofensivos y no ofensivos

**Textos ofensivos en Jovenesweb**


Buscar

Contenido	↑↓ Ofensivo ↓↑
Blanditiis cupiditate ex quia dolore porro.	No ofensivo
Consequatur nihil ut ut dolorum saepe.	Ofensivo
Delectus esse corporis quia quo aut sed.	No ofensivo
Dolores ab saepe repellendus recusandae in.	No ofensivo
Dolores error a consectetur.	Ofensivo
Eligendi a voluptas autem suscipit odit.	Ofensivo
Et sapiente voluptate ullam.	Ofensivo
Ex natus amet incidunt nam.	Ofensivo
Exercitationem non aut officiis nobis et doloribus consequatur.	Ofensivo
Explicabo illo laboriosam fugiat quis.	No ofensivo

Mostrar página 1 de 2  
 Mostrar  por página

[Anterios](#)
[1](#)
[2](#)
[Siguiete](#)

*Nota.* El gráfico representa el resultado de la categorización del contenido

En la Figura 14, se puede identificar como el resultado de la categorización del contenido, el cual se muestra a través de un listado donde se muestra si el texto analizado es ofensivo o no, además está disponible un cuadro de búsqueda donde el cliente puede encontrar la información de manera más rápida, también puede seleccionar cada columna para ordenar los datos de forma ascendente o descendente, según sea el caso y finalmente se agregó la paginación de la información.

**Figura 15***Categorización de las entidades*

The screenshot shows the ContentModerator application interface. On the left is a dark sidebar menu with the following items: Inicio, Crear grupo, Grupos (highlighted), Jovenesweb, Resumen, Textos, Textos ofensivos, Entidades (expanded), Categorías, Información Personal, Enlaces, and menu.pages. The main content area is titled 'Textos analizados en Jovenesweb' and contains a table with the following data:

Id	contenido	Contenido
27		La respuesta esta mal!! yo la resporte p respuest...
78		Es: Descanso es a trabajo La respuesta
124		eso esta mal por que la respuesta en las tien...
211		Ya entiendo ahora si concuerdo con la r
227		Bueno, eso es verdad buena respuesta

*Nota.* El gráfico representa la categorización de las entidades

En la Figura 15, en el menú lateral izquierdo se muestra la categorización de las entidades, la cual se dividió en tres secciones Categorías, Información Personal y Enlaces, en cada sección se agregó un cuadro de búsqueda para filtrar la información, además se puede ordenar el contenido de cada columna de forma ascendente o descendente y también se paginó el contenido.

Figura 16

## Documentación de las APIs disponibles

The image shows a screenshot of an API documentation page for 'Analizar Entidades'. On the left is a dark sidebar with a search bar and a list of navigation items including 'Introduction', 'Authenticating requests', 'Análisis de Entidades', and 'Contenido ofensivo'. The main content area is light blue and contains the following information:

- Analizar Entidades** (requires authentication)
- Esta URL te permite analizar las entidades existentes en un texto
- Se retorna un objeto con el análisis del texto
- Request** Try it out +
- POST** `api/analyze/text/entity`
- Body Parameters**
- `text` string: Texto a ser analizado
- `text_id` integer *optional*: Identificador del texto analizado

On the right, a code editor shows a JavaScript example request:

```
bash javascript
Example request:
const url = new URL(
  "http://content-moderator.test/api/analyze/text/entity"
);
const headers = {
  "Content-Type": "application/json",
  "Accept": "application/json",
};
let body = {
  "text": "nulla",
  "text_id": 14
}
fetch(url, {
  method: "POST",
  headers,
  body: JSON.stringify(body),
}).then(response => response.json());
```

*Nota.* El gráfico representa la documentación de las APIs

En la Figura 16, se puede evidenciar las APIs que están disponibles en Content Moderator, se documentó cada API con una descripción, información de los parámetros de cada URL, detalles del tipo de petición (POST, GET, etc...) necesario para realizar la consulta y un ejemplo en código de cómo realizar la petición con los parámetros requeridos.

### **Entregable Sprint 3**

El Sprint 3 con duración de 2 semanas se realizó del 30/05/2022 hasta el 27/06/2022, en el cual se desarrollaron distintas tareas para dar cumplimiento a las historias de usuario HU09, HU10, y HU11, de las cuales a continuación se mostrarán los avances realizados:

## Figura 17

### Generación de reportes en PDF de cada identidad

Content Moderator				
id contenido	Contenido	Tipo	Sentimiento	Estado
2	Ponen la pregunta mas difícil justo, la que no hay tutoriales de youtube ni en Google. Ni Juan ni AL...	-	N/A	success
12	Ponen la pregunta mas difícil justo, la que no hay tutoriales de youtube ni en Google	-	N/A	success
19	Que falta de respeto con esa grafía cancerígena caballero.	comment	N/A	success
27	La respuesta esta mal! yo la respote pero no para que cambien el orden de las opciones La respuest...	comment	N/A	success
29	Tienes razón la solución ha sido borrada por falta de grafía	comment	N/A	success
33	153 .....85 x.....100 x= 180 180...80 x.....100 x= 225 225.....	comment	N/A	success
46	La excreci n es un proceso fisiol gico, que le permite al organismo eliminar sustancias de desecho L...	solution	neutral	success
49	Una pregunta porque al 80 le resta la edad de Ana ?	comment	N/A	success
50	$x-25x/100-15x/100=975$ $100x-25x-15x/100=975$ $60x/100=975$ $3x/5=975$ $3x=(5)975$ $x=4875/3$ $x=1625//$	comment	N/A	success
56	Le restamos 6 cm al ancho y al largo de las medidas externas del porta retrato porque son 3 cm que o...	comment	N/A	success
57	se me olvidó mensionas que 120 es mayor q 100 por lo tanto se le puede aser el descuento q ofresia e...	comment	N/A	success
72	Tienes razón sergio la respuesta ha sido corregida	comment	positive	success
76	La respuesta es conclusivo porque el parrafo subrayado con negrita comienza con un conector de concl...	comment	N/A	success
78	Es: Descanso es a trabajo La respuesta ya ha sido corregida.	comment	neutral	success
79	esta bien? no deberia ser melodía:oido	comment	N/A	success
84	N se de donde sacas esa solución pues a mí me salió 1024 y no fue la correcta no comprendo	comment	negative	success
85	En la solución esta como se resuelve Y como sacaron 1024?	comment	neutral	success
86	por que? si me dicen razón no se trata d una progresión aritmética? si es como nos están indicando e...	comment	N/A	success
92	jajaj listo, esta esta mejor, la solución es $2^{(10)}$	comment	N/A	success
100	$x + 15/100x = 207$ $100x + 15x/100 = 207$ $115x = 207 \cdot 100$ $x = 20700/115$ $x = 180$	comment	N/A	success
101	Como aclaración hay que tener en cuenta que dos cantidades pueden compararse de dos maneras: Hallan...	comment	N/A	success
114	2- Medidas de superficie Sirven para medir superficies cuadradas, es decir, en dos dimensiones: larg...	comment	N/A	success
119	Se nota que hay triángulos, esferas y cuadrados: el cuadrado gira en diagonal y vuelve a su posició...	comment	neutral	success
124	eso esta mal por que la respuesta en las dos ultimas opciones es la misma disacárido de glucosa tien...	comment	N/A	success

*Nota.* El gráfico representa los reportes en PDF por cada entidad

En la figura 17, la generación de reportes en formato PDF por cada una de las entidades (entidades de categoría, enlaces, información personal) y contenido ofensivo, los reportes son generados a partir de un botón en la parte superior con el icono de PDF, el reporte se genera de manera independiente por entidad, el archivo resultante refleja una fiel copia de la información visualizada en la pantalla.

Figura 18

Generación de los reportes en Excel de cada entidad

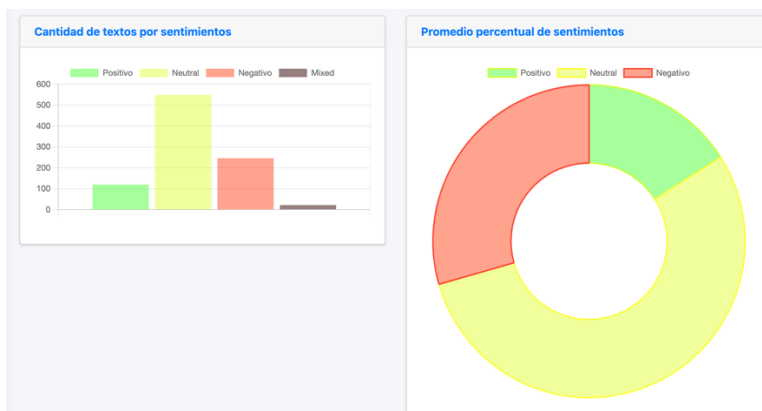
	A	B	C
1	Content Moderator		
2	Texto	Total Entidades	Categorías
3	"Manso" significa que tiene una condición suave, benigna, no bravo; p		1 Quantity
4	3x-3-5 6 Primero lo que esta dividiendo se pasa a multiplicar con un		10 Quantity
5	¿No fue Eloy Alfaro el que firmó el contrato Harman para la construc		3 Person
6	¿Cómo se responde esta pregunta? Para responder correctamente es		24 Skill
7	¿Cómo se responde esta pregunta? Para responder correctamente es		26 Skill
8	Yo tengo otra idea de verlo se repite 3 veces 1,luego 2 y así sucesivam		3 Quantity
9	yo también me salio en regla de 3, dude de mi respuesta		1 Quantity
10	yo pienso que se puede hacer por regla de 3 si en 10 dias se supone q		8 Quantity
11	Yo pienso que es 4, ya que las demas estan hablando de sentimientos		1 Quantity
12	yo no quiero ser arquitecto conchetu quiero ser chefsito		3 PersonType
13	Yo no entendi nada :/ y mañana me toman el examen :(		1 DateTime
14	yo me guie por las flechas pequeñas , girandolas como manisillas de r		2 Product
15	Yo lo hice dividiendo 1 boli al azar entre los bolis que contiene en su e		4 Quantity
16	yo llegue a la conclusion de que tuvo = tener tubo = es un tubo algo ci		2 Product
17	Yo le hice con gráfica, dibujo un rectangulo dividido en 5 partes de esa		8 Skill
18	Yo fui multiplicando por 2. 4x2=8 8x2=16		3 Quantity
19	Yo creo que esa no es la respuesta ya que igual falta una línea, el trián		4 Quantity
20	yo enteedy k yo see multiplica 5+9=45 5.28= 140 el porque si noc		5 Quantity
21	ya le entendi solo ha sido de multiplicar x 2 nada mas esto si es facil.		1 Quantity
22	Ya entiendo ahora si concuerdo con la respuesta 18-13=234+1300=1!		5 DateTime
23	ya entendi... es la diferencia del segundo cuadro menos el primero, ese		5 Quantity

*Nota.* El gráfico representa los reportes en Excel de cada una de las entidades

En la figura 18, la generación de reportes en formato Excel por cada una de las entidades (entidades de categoría, enlaces, información personal) y contenido ofensivo, los reportes son generados a partir de un botón verde que se encuentra en la parte superior izquierda con el icono de Excel, el reporte se genera de manera independiente por entidad, el archivo resultante refleja una fiel copia de la información visualizada en la pantalla.

## Figura 19

### Generación de gráficos estadísticos



*Nota.* El gráfico representa los gráficos estadísticos de del análisis del contenido

En la figura 19, se observa la generación de gráficos estadísticos generados a partir del análisis del contenido, el primer gráfico podemos observar la cantidad de sentimientos, mientras que en el segundo gráfico es la representación porcentual de los sentimientos con el sentimiento predominante, los colores se estandarizan para lograr mayor facilidad al momento de interpretar los gráficos.

### **Entregable Sprint 4**

El Sprint 4 con duración de 2 semana se realizó del 13/06/2022 hasta el 24/06/2022, en el cual se desarrollaron distintas tareas para dar cumplimiento a las historias de usuario HU11, HU12, HU13 y HU04, de las cuales a continuación se mostrarán los avances realizados:

## Figura 20

### Porcentaje de contenido ofensivo y palabras más repetidas



*Nota.* El gráfico representa las gráficas de pastel y barras de la información analizada

En la figura 20, observamos el desarrollo de las gráficas de pastel y barras cada una representado la información del contenido analizado, en el primer gráfico observamos el porcentaje de contenido ofensivo y no ofensivo, mientras en la segunda grafica de barras observamos las palabras ofensivas más repetidas.

## Figura 21

### Gráfico estadístico de barras con las entidades más repetidas



*Nota.* El gráfico representa las gráficas estadísticas de las entidades

En la figura 21, observamos el gráfico estadístico generado a partir del análisis del contenido, se puede observar claramente las diferentes entidades y la cantidad que existe en cada una de ellas, para el mejor entendimiento del usuario, se realiza un gráfico de barras, la gráfica se realiza de color azul como estandarización.

## Figura 22

*Tablas dinámicas para el manejo de información*

Entidades identificadas en Jovenesweb

Buscar

Texto	Total Entidades	Categorías	Acción
$A = b \cdot h$ $6x^2 + 22x + 20 = (3x+5) \cdot (?????)$ Multiplicamos solo los términos independientes $20 = 5 \cdot ?$ $20 = 5 \cdot 4$ el +4 es el único que cumple que respuestas tiene +4?? la B) y la C). Multiplicamos con la base y tiene que dar el área Se aplica la propiedad distributiva $\downarrow B) (3x+5)(3x+4) = 9x^2 + 12x + 15x + 20 = 9x^2 + 27x + 20$ (No es el área)C) $(3x+5)(2x+4) = 6x^2 + 12x + 10x + 20 = 6x^2 + 22x + 20$ (SI ES EL ÁREA) Respuesta C)	4	<ul style="list-style-type: none"> <li>Quantity</li> <li>Quantity</li> <li>Skill</li> <li>Quantity</li> </ul>	Ver
$45 \text{ Sen} \theta = 45 \sqrt{3} \cdot \text{Sen } 30^\circ$ (El 45 pasa a dividir del otro lado, así que eso igual se lo elimina) $\text{Sen} \theta = \sqrt{3} \cdot 1/2 \Leftarrow$ (Esto sale del valor de $\text{Sen } 30^\circ$ ) $\text{Sen} \theta = \sqrt{3}/2 = 60^\circ$ Ya que al ubicarnos en la tabla de ángulos, el $\sqrt{3}/2$ se encuentra en el $\text{Sen } 60^\circ$ )	7	<ul style="list-style-type: none"> <li>Quantity</li> <li>Quantity</li> <li>Quantity</li> <li>Quantity</li> <li>...</li> </ul>	Ver
$(C^3 C^2)$ La que es con la misma base que es "C" La regla es: cuando tiene la misma base se suman los exponentes $3+2=5$ Y el 5 lo multiplicamos por el 2 del exponente de afuera que viene hacer 10 La respuesta es $C^{10}$	5	<ul style="list-style-type: none"> <li>Quantity</li> <li>Quantity</li> <li>DateTime</li> <li>Quantity</li> <li>...</li> </ul>	Ver

*Nota.* El gráfico representa las tablas dinámicas permitiendo el manejo de información de eficiente

En la figura 22, el uso de tablas dinámicas nos permite el manejo de la información de una forma más libre y optima de organizar, desde el ordenamiento de la información en base a una columna e incluso filtrarla según las necesidades del usuario, se encuentra desarrollado por cada una de las categorías, contenido ofensivo y por los textos analizados.

## Etapa 5: Implementación

La implementación del aplicativo web Content Moderator, se desarrolló de forma iterativa en conjunto con el equipo de Jovenesweb usando las siguientes herramientas:



**GitHub:** Es un repositorio digital que nos permite almacenar el código para el control de versionamiento y el trabajo en equipo teniendo acceso de manera remota, el proyecto incluye una licencia y el archivo de README el cual es información acerca del proyecto (GitHub, 2022).

## Figura 23

### Repositoria del proyecto

The screenshot shows the GitHub interface for the repository 'insoutt/content-moderator'. The repository is private and has 1 branch (main), 0 tags, and 72 commits. The commit history is listed below, showing files and folders added or modified over time. The right sidebar provides information about the repository, including a link to the website 'moderator.betaksur.com', and a 'Languages' section showing the code is primarily PHP (62.5%) and Blade (37.2%), with a small amount of Shell (0.3%).

File/Folder	Commit Message	Time Ago
app	Usar Https	34 minutes ago
bootstrap	Primer commit	12 months ago
config	Quitar enlace	7 days ago
database	Agregar estado de contenido analizado	last month
public	colors	10 months ago
resources	Refactorizar	7 days ago
routes	textos ofensivos	7 days ago
storage	Primer commit	12 months ago
tests	Primer commit	12 months ago
.editorconfig	Primer commit	12 months ago
.env.example	Primer commit	12 months ago
.gitattributes	Primer commit	12 months ago
.gitignore	Implementar documentación API	11 months ago
.styleci.yml	Primer commit	12 months ago
README.md	Actualizar documentación	11 months ago
artisan	Primer commit	12 months ago
composer.json	Implementar documentación API	11 months ago
composer.lock	Implementar documentación API	11 months ago
package-lock.json	Primer commit	12 months ago
package.json	Primer commit	12 months ago

**Languages**

- PHP 62.5%
- Blade 37.2%
- Shell 0.3%

*Nota.* El gráfico representa el repositorio del proyecto almacenado en GitHub.

Para el desarrollo de Content Moderator, se optó por crear un repositorio en GitHub para mantener un control de cada cambio realizado en el proyecto, esto facilitó el trabajo en equipo ya que cada miembro del equipo SCRUM de desarrollo pudo trabajar de manera conjunta manteniendo siempre los últimos cambios del equipo.

**DigitalOcean:** Es un proveedor de diferentes servicios, nos permite de manera sencilla la creación de servidores para el despliegue de la aplicación, mediante la virtualización para la optimización de los recursos (DigitalOcean, 2022).

**Forge:** Es un administrador de los diferentes servidores que puede necesitar una aplicación tales como: servidor web, servidor de trabajo, servidor de cache o servidor de base de datos, esta herramienta permite el despliegue de aplicaciones web desarrolladas en PHP y en sus diferentes frameworks (FORGE, 2022).

## Figura 24

*Servidor que almacena el repositorio*

The screenshot shows the Forge web interface for a site named 'moderator. ur.com'. The interface includes a navigation menu on the left with options like App, Deployments, Commands, Packages, Notifications, Environment, Queue, SSL, Security, Redirects, and Logs. The main content area displays the 'Live Deployment Output' and a 'Deployment History' table. The table lists the following deployment records:

When	Initiated By	Branch	Commit	Duration	Status
4 hours ago	EB	main	b136d86c	12 seconds	Finished
4 hours ago	EB	main	0dafa5f5	16 seconds	Finished
1 month ago	EB	main	e38adfa9	9 seconds	Finished
1 month ago	EB	main	e38adfa9	—	Failed

Showing the last 10 deployments. Deploying the branch.

*Nota.* El gráfico el servicio para desplegar el proyecto de Laravel en un servidor

Debido a que Content Moderator fue desarrollado en Laravel y gracias al apoyo de Jovenesweb se optó por usar Forge para poder realizar el deploy del aplicativo web. Esta herramienta permitió la creación y configuración del ambiente de desarrollo para el proyecto de manera automática además permite mantener actualizado el servidor con los últimos cambios que se realizan en la rama principal del repositorio alojado en GitHub.

## Capítulo IV

### Validación de la propuesta

#### Introducción del capítulo

Con la finalización del desarrollo del aplicativo web para el análisis de contenido usando servicios cognitivos de Inteligencia Artificial, se procede a realizar una validación en conjunto con el equipo de Jovenesweb, el desarrollo de la validación es bajo un carácter práctico en el cual se validarán cada uno de los indicadores establecidos, el mismo que tiene como objetivo demostrar la optimización de los tiempos y calidad de revisión de contenido en la plataforma Jovenesweb. mediante la recolección de datos de los resultados de cada indicador establecido.

Para este fin, se implementará un análisis estadístico de validación en cada uno de los indicadores definidos, para la realización de un análisis comparativo los datos son tomados cuando el análisis de contenido se realizaba de manera manual y al realizarlo con el aplicativo web implementado.

#### Obtención de datos

Los datos obtenidos para el análisis estadístico son de fuentes primarias, en este caso provienen directamente del análisis de contenido de la plataforma educativa Jovenesweb, los mismos que no han sido analizados, ni procesados en ninguna otra investigación y es información de primera mano, los datos tomados son de tipo cuantitativo, de esta manera permite demostrar si la hipótesis nula o alternativa se cumple por cada indicador.

Los datos para los indicadores del 1 al 4, provienen de los registros históricos de Jovenesweb, los mismos que habían sido analizados para entender en comportamiento de los usuarios en el aplicativo estos datos han sido validados por parte de Jovenesweb.

Además, para la validación se ha tomado la información que el aplicativo genera y de esta manera determinar si cumple con la hipótesis planteada por cada indicador, en los indicadores que solicitan el porcentaje para la demostración se utilizara la cantidad para de esta manera.

Se debe recalcar que el porcentaje es calculado en base a la cantidad por ende no existe limitaciones para considerarlo, mientras que los dos últimos indicadores (indicador 5 y 6) no existe un registro histórico, sin embargo el equipo de Jovenesweb a realizada una revisión de manera exhaustiva el mismo contenido analizado con el objetivo de compararlo y demostrar el cumplimiento de la hipótesis, se debe considerar que por políticas de la empresa Jovenesweb se tuvo un acceso a una muestra de mil textos generados por el usuario de la plataforma Jovenesweb.

### ***Datos recolectados***

**Tabla 35**

*Frecuencia esperada y frecuencia observada del indicador 1*

	Frecuencia esperada Jovenesweb	Frecuencia observada Content Moderator
Positivo	16,28	15,96
Negativo	33,24	29,45
Neutral	50,48	54,62

**Tabla 36**

*Frecuencia esperada y frecuencia observada del indicador 2*

	Frecuencia esperada Jovenesweb	Frecuencia observada Content Moderator
% Ofensivo	9,4	0,77
% No ofensivo	90,6	99,23

**Tabla 37***Frecuencia esperada y frecuencia observada del indicador 3*

	Frecuencia esperada Jovenesweb	Frecuencia observada Content Moderator
Tiempo análisis sentimientos	810,24	0,0319444
Tiempo análisis ofensivo	810,24	0,0333333
Tiempo análisis entidades	810,24	0,0655556

**Tabla 38***Frecuencia esperada y frecuencia observada del indicador 4*

	Frecuencia esperada Jovenesweb	Frecuencia observada Content Moderator
Porcentaje de contenido analizado de manera exitosa	90	90
Porcentaje de error de contenido analizado	10	10

**Tabla 39***Frecuencia esperada y frecuencia observada del indicador 5*

	Frecuencia esperada Jovenesweb	Frecuencia observada Content Moderator
Cantidad Categoría	2817	3626
Cantidad Información personal	356	527
Cantidad enlaces	597	692

**Tabla 40***Frecuencia esperada y frecuencia observada del indicador 6*

	Frecuencia esperada Jovenesweb	Frecuencia observada Content Moderator
Positivo	132	120
Negativo	219	246
Neutral	564	549

## **Análisis de datos**

Por medio del de distribución de probabilidad estadística chi-cuadrado, el cual nos permite validar las hipótesis entre dos variables a través de la distribución de la frecuencia, como primer paso se usará el estadístico chi-cuadrado para después se evaluará la distribución obtenida con la distribución esperada o teórica.

El cálculo del valor crítico se realiza de manera independiente para cada indicador, para lo cual también se considerará el cálculo de los grados de libertad con la siguiente fórmula:  $gl = (\text{num columnas} - 1) * (\text{num filas} - 1)$ , el nivel de significancia será de 5% ( $\alpha = 0,05$ ).

Para la validación de los indicadores del 1 al 6, se utilizará la prueba de bondad del chi cuadrado dada la naturaleza de sus variables categóricas, la misma que nos permite evaluar la hipótesis de una manera estadística desde una distribución no específica, por ende, los datos son representativos desde una perspectiva de la población completa (Machines, 2021).

### ***Validación para el indicador 1***

Formulación de hipótesis para el indicador: porcentaje de sentimientos positivos, negativos y neutrales de los textos analizados

- **H0:** El aplicativo web se ajusta a la distribución teórica del porcentaje de contenido analizado en los sentimientos: positivos, negativos y neutrales
- **H1:** El aplicativo web no se ajusta a la distribución teórica de porcentaje de contenido analizado en los sentimientos: positivos, negativos y neutrales

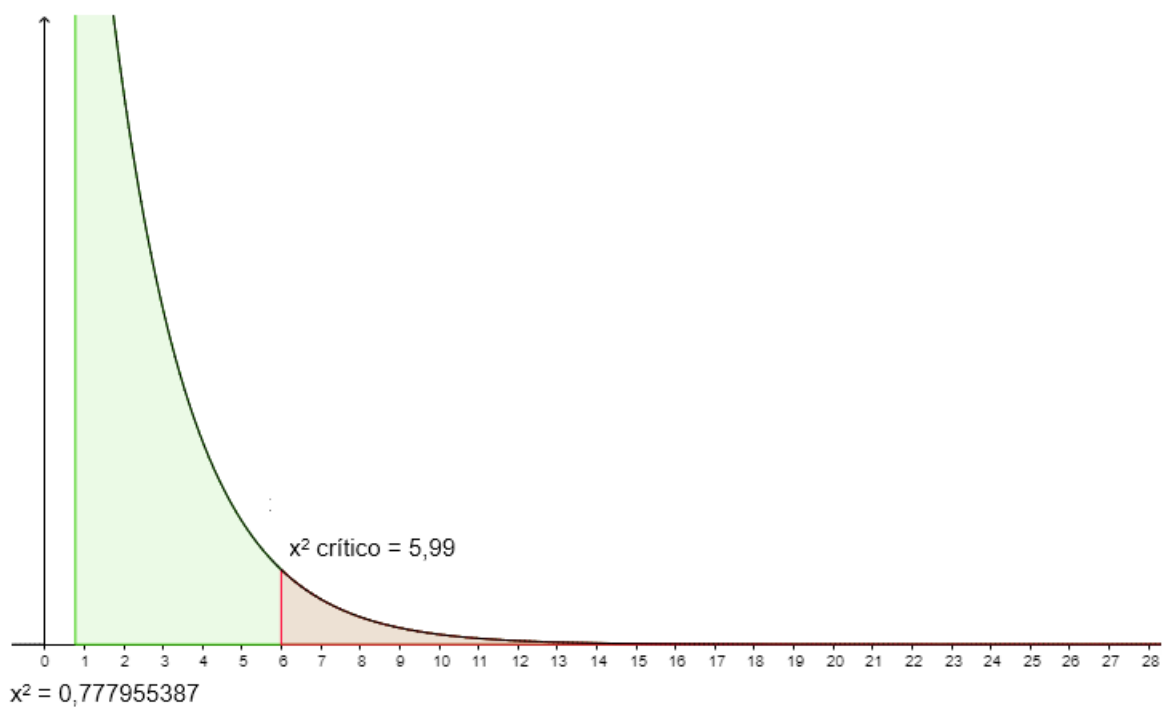
**Tabla 41**

*Calculo de bondad de ajuste de chi-cuadrado del indicador 1*

	Frecuencia esperada Jovenesweb	Frecuencia observada Content Moderator	Total
Positivo	16,28	15,96	0,006289926
Negativo	33,24	29,45	0,432132972
Neutral	50,48	54,62	0,339532488
		Chi-cuadrado calculado	0,777955387

**Figura 25**

*Gráfico de chi-cuadrado del indicador 1*



Se acepta la hipótesis nula (H0) al ser el valor chi-cuadrado calculado menor al valor crítico, se rechaza la hipótesis alternativa (H1).



### **Validación para el indicador 2**

Formulación de hipótesis para el indicador: porcentaje de contenido ofensivo detectado

- **H0:** El aplicativo web se ajusta a la distribución teórica del porcentaje de contenido ofensivo detectado
- **H1:** El aplicativo web no se ajusta a la distribución teórica del porcentaje de contenido ofensivo detectado

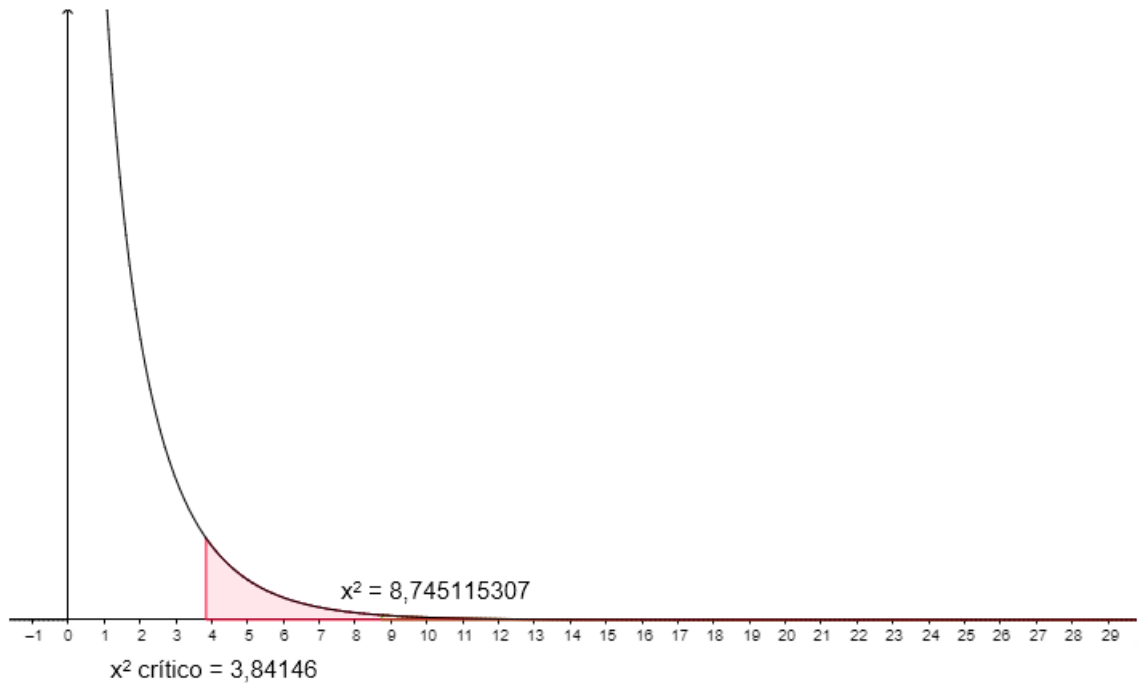
### **Tabla 42**

*Calculo de bondad de ajuste de chi-cuadrado del indicador 2*

	Frecuencia esperada	Frecuencia observada	
	Jovenesweb	Content Moderator	Total
% Ofensivo	9,4	0,77	7,923074468
% No ofensivo	90,6	99,23	0,822040839
		Chi-cuadrado calculado	8,745115307

**Figura 26**

*Gráfico de chi-cuadrado del indicador 2*



Se rechaza la hipótesis nula ( $H_0$ ) al ser el valor chi-cuadrado calculado mayor al valor crítico, se acepta la hipótesis alternativa ( $H_1$ ).

### ***Validación para el indicador 3***

Formulación de hipótesis para el indicador: tiempo en el análisis de contenido

- **H0:** El tiempo de análisis del aplicativo web es mayor a al tiempo de análisis de Jovenesweb
- **H1:** El tiempo de análisis del aplicativo web es menor a al tiempo de análisis de Jovenesweb

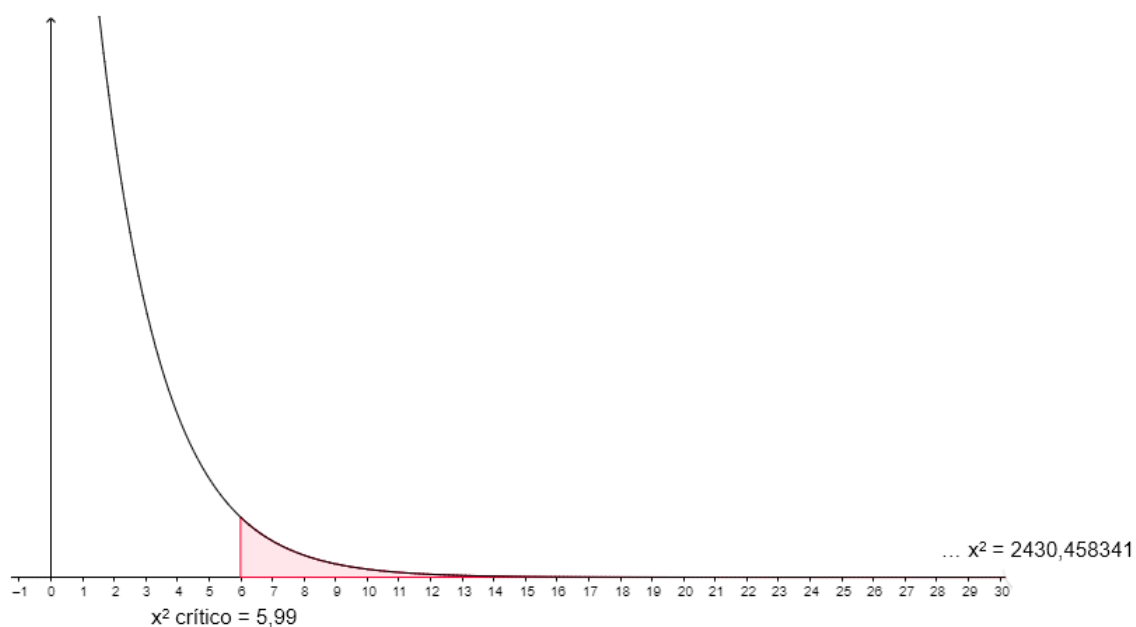
**Tabla 43**

*Calculo de bondad de ajuste de chi-cuadrado del indicador 3*

	Frecuencia esperada Jovenesweb	Frecuencia observada Content Moderator	Total
Tiempo análisis sentimientos	810,24	0,0319444	810,1761125
Tiempo análisis ofensivo	810,24	0,0333333	810,1733348
Tiempo análisis entidades	810,24	0,0655556	810,1088941
		Chi-cuadrado calculado	2430,458341

**Figura 27**

*Gráfico de chi-cuadrado del indicador 3*



Se rechaza la hipótesis nula ( $H_0$ ) al ser el valor chi-cuadrado calculado mayor al valor crítico, se acepta la hipótesis alternativa ( $H_1$ ).

### Validación para el indicador 4

Formulación de hipótesis para el indicador: porcentaje del contenido analizado

- **H0:** El aplicativo web se ajusta a la distribución teórica en el porcentaje de contenido analizado.
- **H1:** El aplicativo web no se ajusta a la distribución teórica en el porcentaje de contenido analizado

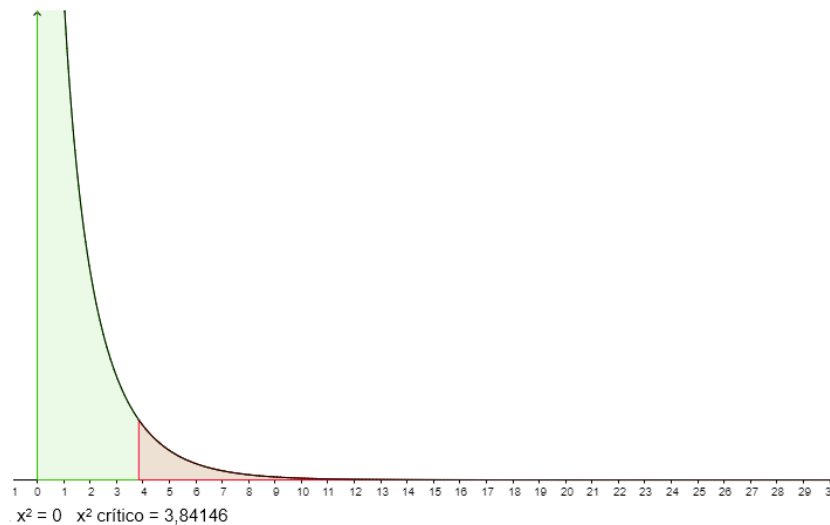
**Tabla 44**

*Calculo de bondad de ajuste de chi-cuadrado del indicador 4*

	Frecuencia esperada Jovenesweb	Frecuencia observada Content Moderator	Total
Porcentaje de contenido analizado de manera exitosa	90	90	0
Porcentaje de error de contenido analizado	10	10	0
		Chi-cuadrado calculado	0

**Figura 28**

*Gráfico de chi-cuadrado del indicador 4*



Se acepta la hipótesis nula ( $H_0$ ) al ser el valor chi-cuadrado calculado menor al valor crítico, se rechaza la hipótesis alternativa ( $H_1$ ).

### **Validación para el indicador 5**

Formulación de hipótesis para el indicador: cantidad de entidades detectadas y clasificadas

- **H0:** El aplicativo web se ajusta a la distribución teórica en las cantidades de entidades detectadas y las clasifica en: categoría, información personal, cantidad de enlaces
- **H1:** El aplicativo web no permite mejorar las cantidades de entidades detectadas y las clasifica en: categoría, información personal, cantidad de enlaces

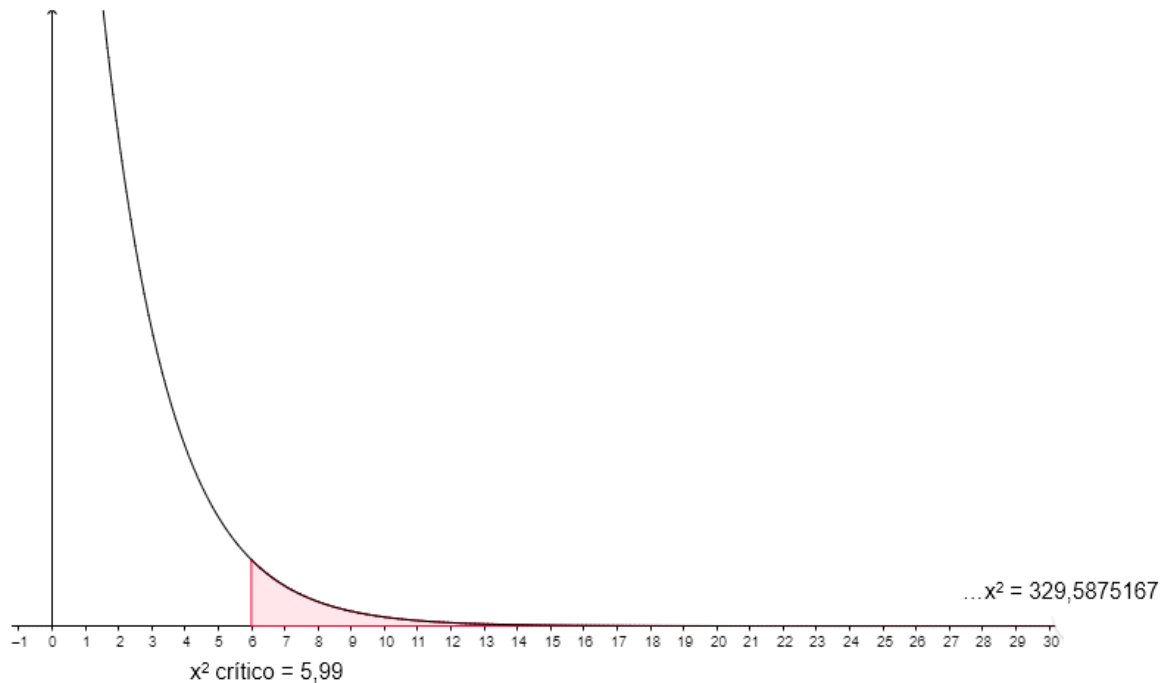
**Tabla 45**

*Calculo de bondad de ajuste de chi-cuadrado del indicador 5*

	Frecuencia esperada Jovenesweb	Frecuencia observada Content Moderator	Total
Cantidad Categoría	2817	3626	232,3326234
Cantidad Información personal	356	527	82,13764045
Cantidad enlaces	597	692	15,11725293
		Chi-cuadrado calculado	329,5875167

## Figura 29

Gráfico de chi-cuadrado del indicador 5



Se rechaza la hipótesis nula ( $H_0$ ) al ser el valor chi-cuadrado calculado mayor al valor crítico, se acepta la hipótesis alternativa ( $H_1$ ).

### Validación para el indicador 6

Formulación de hipótesis para el indicador: Cantidad de contenido con sentimientos identificados y clasificado (positivo, negativo y neutro)

- **H0:** El aplicativo web se ajusta a la distribución teórica en las cantidades de sentimientos detectadas y los clasifica en: positivo, negativo y neutral
- **H1:** El uso del aplicativo web no permite mejorar las cantidades de sentimientos detectadas y los clasifica en: positivo, negativo y neutral

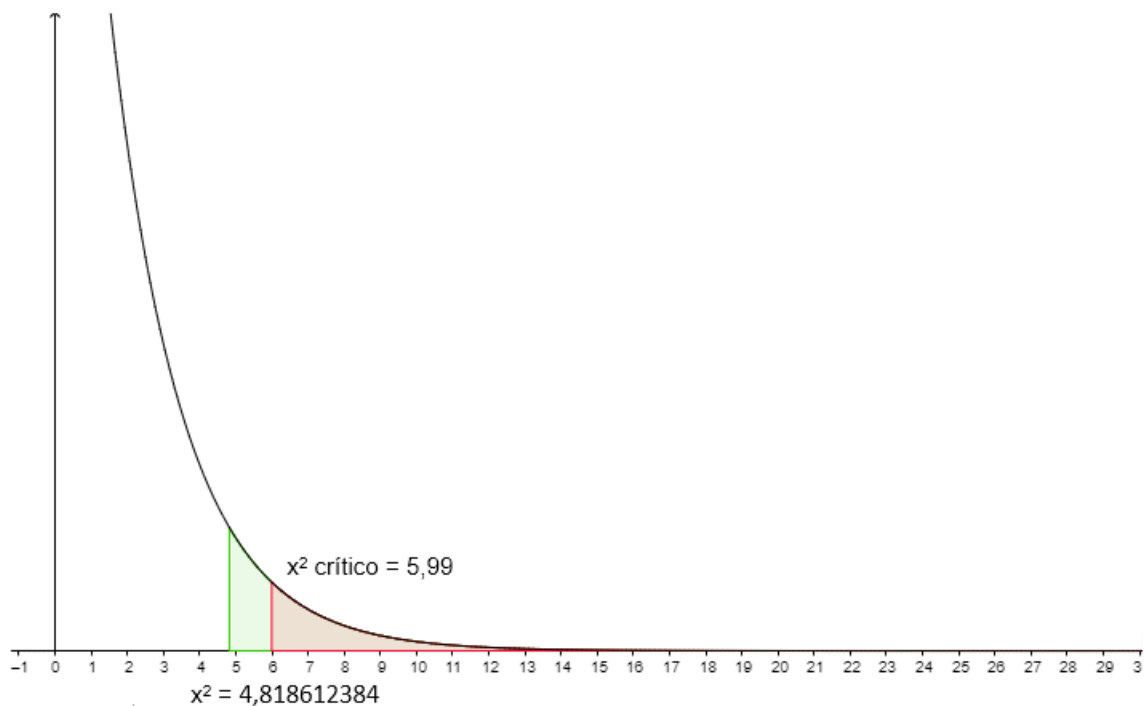
**Tabla 46**

*Calculo de bondad de ajuste de chi-cuadrado del indicador 6*

	Frecuencia esperada Jovenesweb	Frecuencia observada Content Moderator	Total
Positivo	132	120	1,090909091
Negativo	219	246	3,328767123
Neutral	564	549	0,39893617
		Chi-cuadrado calculado	4,818612384

**Figura 30**

Gráfico de chi-cuadrado del indicador 6



Se acepta la hipótesis nula ( $H_0$ ) al ser el valor chi-cuadrado calculado menor al valor crítico, se rechaza la hipótesis alternativa ( $H_1$ ).

### **Presentación de resultados del análisis de los datos**

Después de realizar el análisis y la toma de decisión acerca en cada criterio con su respectiva hipótesis y se presenta los siguientes resultados:

El aplicativo web se ajusta a la distribución teórica del porcentaje de contenido analizado en los sentimientos: positivos, negativos y neutrales

El aplicativo web no se ajusta a la distribución teórica del porcentaje de contenido ofensivo detectado

El tiempo de análisis del aplicativo web es menor a al tiempo de análisis de Jovenesweb

El aplicativo web se ajusta a la distribución teórica en el porcentaje de contenido analizado.

El aplicativo web no permite mejorar las cantidades de entidades detectadas y las clasifica en: categoría, información personal, cantidad de enlaces

El aplicativo web se ajusta a la distribución teórica en las cantidades de sentimientos detectadas y los clasifica en: positivo, negativo y neutral



## Capítulo V

### Conclusiones y recomendaciones

#### Conclusiones

- Se cumplió con el objetivo de desarrollar un aplicativo web que permita optimizar el análisis de contenido publicado por los usuarios de la plataforma web educativa Jovenesweb con el uso de servicios cognitivos de Inteligencia artificial de Microsoft Azure
- La elaboración del marco teórico permitió el entendimiento de los conceptos fundamentales para el desarrollo de un aplicativo web que permita implementar los servicios cognitivos de Inteligencia Artificial ofrecidos por Microsoft Azure para realizar el reconocimiento de: entidades, sentimientos y contenido ofensivo a fin de automatizar el moderamiento de contenido en la plataforma Jovenesweb
- A través de la integración del aplicativo web con los servicios cognitivos de Inteligencia Artificial de Microsoft Azure optimizó el análisis del contenido generado por los usuarios de Jovenesweb.
- El desarrollo del aplicativo web por medio del marco de trabajo Scrum permitió obtener un dinamismo de trabajo sostenible, priorizando las funcionalidades por su valor agregado, estimando el esfuerzo requerido en cada una de ellas y respondiendo con agilidad a los cambios.
- Para el desarrollo del aplicativo web se utilizó el framework Laravel 8 de PHP en conjunto con el servicio de Laravel Forge, el cual permite desplegar y configurar el aplicativo en un servidor de DigitalOcean para que sea accesible a través de internet
- Se implementó de manera satisfactoria el aplicativo web en conjunto con el equipo de desarrollo de Jovenesweb y en base al cronograma de entregas establecido en la etapa del inicio del proyecto

- Para la validación de los indicadores establecidos se utilizó la técnica estadística del Chi cuadrado por medio de bondad de ajuste con un nivel de significancia del 5% lo que comprobó la hipótesis.
- Se debe tener en cuenta que el análisis de contenido usando los servicios cognitivos de Inteligencia Artificial de Microsoft Azure está sujeto a limitaciones que podrían provocar resultados no deseados como; el mal uso de las reglas gramaticales el cual distorsiona el contexto de la oración y el uso de abreviaturas que dificulta el reconocimiento de contenido ofensivo e identidades

### **Recomendaciones**

- Se recomienda trabajar con la versión de pago de los servicios cognitivos de Inteligencia Artificial de Microsoft Azure, ya que en la versión gratuita existen limitaciones sobre la cantidad de textos a analizar al mes y limitaciones en la cantidad de textos a analizar por minuto los cuales podrían aumentar el tiempo de análisis.
- Se recomienda el uso del marco de trabajo SCRUM para facilitar el desarrollo del proyecto dada su flexibilidad y adaptación.
- Se recomienda antes de analizar el contenido validar que las reglas gramaticales sean las correctas y que no existen abreviaturas en palabras para evitar resultados inesperados en el análisis.
- Se recomienda normalizar los textos antes de ser analizados, eliminando código HTML, letras duplicadas, espacios en blanco duplicados para evitar resultados inesperados en el análisis de entidades, sentimientos o contenido ofensivo mediante la aplicación de expresiones regulares y funciones de texto propias del lenguaje de programación.
- Se sugiere leer la documentación de los servicios cognitivos de Inteligencia Artificial de Microsoft Azure para poder comprender el funcionamiento de la API y sus limitaciones.

- Para futuras investigaciones con temáticas similares se recomienda integrar el análisis de moderación de contenido para imágenes cuya integración es similar a la moderación de contenido ofensivo en textos y que permitirá evaluar la eficacia del análisis.

## Bibliografía

- Alejandra, C. (2019). *Repositorio institucional de la UNLP*. Obtenido de Minería de opiniones: análisis de sentimientos en una red social: <http://sedici.unlp.edu.ar/handle/10915/77379>
- Alicia, P. (2018). *Revista Iberoamericana de Inteligencia Artificial* . Obtenido de Tecnicas de extraccioon de entidades con nombre: <https://www.redalyc.org/pdf/925/92530455002.pdf>
- Anibal, J. (06 de 08 de 2020). *PROYECTO DE LEY ORGANICA LEY PARA PREVENIR LA VIOLANCIA, EL ACOSO DIGITAL Y LA VIOLACION A LA INTIMIDAD*. Obtenido de <http://www.asobanca.org.ec/sites/default/files/PROYECTO%20DE%20LEY%20ORG%C3%81NICA%20PARA%20PREVENIR%20LA%20VIOLENCIA%2C%20EL%20ACOSO%20DIGITAL%20Y%20LA%20VIOLACI%C3%93N%20A%20LA%20INTIMIDAD.pdf>
- Balaguer, M. L. (16 de 01 de 2019). *Dialnet*. Obtenido de Libertad de expresión y redes sociales: ¿es posible sancionar al trabajador por publicaciones de contenido privado ajenas al trabajo?: <https://dialnet.unirioja.es/servlet/articulo?codigo=6959609>
- Barcelona, F. d. (2018). *Història d'Internet*. Obtenido de <https://www.fib.upc.edu/retroinformatica/historia/internet.html>
- Baviera, T. (2017). *Universitat Pilitecnica de Valencia* . Obtenido de Técnicas para el análisis del sentimiento en Twitter: Aprendizaje Automático Supervisado y SentiStrength: <https://riunet.upv.es/handle/10251/153230>
- Britannica, T. E. (23 de 02 de 2022). *Myspace. Encyclopedia Britannica*. Obtenido de <https://www.britannica.com/topic/Myspace>
- Bruno, T. (11 de 05 de 2017). *El Mundo*. Obtenido de El horroroso trabajo del moderador de contenido online: <https://www.elmundo.es/tecnologia/2017/05/11/590b008546163f04668b4609.html>

Builtin. (2020). *Introduction to IA*. Obtenido de <https://builtin.com/artificial-intelligence>

Calcano Monts, M. A. (2021). *Internet, redes sociales, libertad de expresión y derechos humanos. en derechos humanos. 10 años*. Obtenido de Comisión de Derechos Humanos de la Ciudad de México: [https://cdhcm.org.mx/wp-content/uploads/2021/06/10-años-de-la-reforma-digital\\_FINAL.pdf#page=232](https://cdhcm.org.mx/wp-content/uploads/2021/06/10-años-de-la-reforma-digital_FINAL.pdf#page=232)

CERN. (2019). *Una breve historia de la Web*. Obtenido de La Web ha crecido para revolucionar las comunicaciones en todo el mundo: <https://home.cern/science/computing/birth-web/short-history-web#:~:text=Tim%20Berners%2DLee%2C%20a%20British,and%20institutes%20around%20the%20world>.

certmind. (2020). *The Scrum Map*. CertMind.

*Cognitive Services: API para desarrolladores de inteligencia artificial | Microsoft Azure*. (2021). Obtenido de Azure: <https://azure.microsoft.com/es-es/services/cognitive-services/#features>

*CONSTITUCION DE LA REPUBLICA DEL ECUADOR 2008*. (20 de 10 de 2008). Obtenido de [https://www.oas.org/juridico/pdfs/mesicic4\\_ecu\\_const.pdf](https://www.oas.org/juridico/pdfs/mesicic4_ecu_const.pdf)

DigitalOcean. (2022). *Web Hosting*. Obtenido de <https://www.digitalocean.com/solutions/website-hosting>

Dubiau, L. |. (2018). *Repositorio instituoinal de la UNLP*. Obtenido de Análisis de sentimientos sobre un corpus en español: experimentación con un caso de estudio: <http://sedici.unlp.edu.ar/handle/10915/76148>

FORGE. (2022). *What is Forge?* Obtenido de <https://forge.laravel.com/docs/1.0/introduction.html>

Garcia, M. (25 de 05 de 2020). *¿Qué es el SPRINT REVIEW?* Obtenido de [https://ittude.com.ar/b/scrum/que-es-el-sprint-review/#:~:text=El%20SPRINT%20REVIEW%20\(o%20Revisi%C3%B3n,que%20tuvo%20el%20Product%20Backlog](https://ittude.com.ar/b/scrum/que-es-el-sprint-review/#:~:text=El%20SPRINT%20REVIEW%20(o%20Revisi%C3%B3n,que%20tuvo%20el%20Product%20Backlog).

Garcia, P. (19 de 01 de 2021). *Servicios cognitivos para mejorar la experiencia de usuario*. Obtenido de <https://blogvisionarios.com/impulsa-tu-negocio/potencia-cloud-data-ia/servicios-cognitivos-ux/>

GitHub. (2022). *Documentaion de GitHub*. Obtenido de <https://docs.github.com/es/get-started/quickstart/hello-world>

Guerrero, C. (25 de 02 de 2021). *Hiperderechos*. Obtenido de Políticas de moderación de contenidos: <https://hiperderecho.org/2020/02/politicas-de-moderacion-de-contenidos/>

Hope, C. (06 de 07 de 2019). *Computer Home* . Obtenido de <https://www.computerhope.com/jargon/c/content.htm>

INEGI. (10 de 04 de 2019). *DE LA POBLACIÓN DE 12 A 59 AÑOS USUARIA DE INTERNET, 16.8% HA VIVIDO*. Obtenido de <https://www.inegi.org.mx/contenidos/saladeprensa/boletines/2019/estsociodemo/mociba-2017.pdf>

Kimin, J. (28 de 02 de 2021). *ThpughtCo*. Obtenido de What Is Web Content?: <https://www.thoughtco.com/what-is-web-content-3466787>

Laoyan, S. (05 de 11 de 2021). *La importancia de la planificación de sprints en las metodologías ágiles*. Obtenido de asana: <https://asana.com/es/resources/sprint-planning-meeting>

Lara-Navarra, P., López-Borrull, A., Sánchez-Navarro, J., & Yànez, P. (2018). *Medición de la influencia de usuarios en redes sociales: propuesta SocialEngagement. El profesional de la información (EPI)*. Obtenido de Profesional de la Información:

<http://profesionaldelainformacion.com/contenidos/2018/jul/18.pdf>

Leiner, B. M., Cerf, V. G., Kahn, D. D., & Wolff, S. (s.f.). *Orígenes de Internet*. Obtenido de <https://www.isoc.org/internet/history-internet/brief-history-internet/>

López-de-Ayala, M. C., Vizcaíno-Laorga, R., & Montes-Vozmediano, M. (2020). Hábitos y actitudes de los jóvenes ante las redes sociales: influencia del sexo, edad y clase social. *Profesional de la Información*, 10.

Machines, I. B. (07 de 12 de 2021). *Prueba de chi-cuadrado de bondad de ajuste*. Obtenido de <https://www.ibm.com/docs/es/spss-statistics/beta?topic=tests-chi-square-test-goodness-fit>

Maldonado Berea, G. A., García González, J., & Sampedro-Requena, B. E. (2019). El efecto de las TIC y redes sociales en estudiantes universitarios. *Revista Iberoamericana de Educación a Distancia*.

Martínez Ortega, A. G., & Medina Chicaiza, R. P. (2020). *Tecnologías en la inteligencia artificial para el Marketing: una revisión de la literatura*. Pro Sciences: Revista de Producción, Ciencias e Investigación.

Megias, A. (s.f.). *SERVICIO WEB DE IDENTIFICACIÓN Y CLASIFICACIÓN DE ENTIDADES*. Obtenido de <https://core.ac.uk/download/pdf/288498333.pdf>

Morales Alarcón, C. H. (2017). *Universidad de las Fuerzas Armadas ESPE*. Obtenido de SLR aplicando Text Mining para la adopción de software de Business Intelligence en el área de investigación de la UNACH: <http://repositorio.espe.edu.ec/handle/21000/13139>

RAE. (2020). *Diccionario de la lengua española (Ofender)*. Obtenido de

<https://dle.rae.es/ofender?m=form>

RAE. (2020). *Diccionario de la lengua española (sentimiento)*. Obtenido de

<https://dle.rae.es/sentimiento?m=form>

Ramírez, M. R., Soto, M. S., Moreno, H. R., Rojas, E. M., Millán, N. O., & Cisneros, R. R.

(2019). Metodología SCRUM y desarrollo de Repositorio Digital. *Revista Ibérica De Sistemas e Tecnologías De Informação*.

Rouhiainen, L. (2018). *Inteligencia Artificial*. Madrid: Alienta Editorial.

Scrum, L. (22 de 08 de 2020). *Técnicas que te ayudarán a Priorizar*. Obtenido de

<https://librescrum.com/2020/09/22/tecnicas-de-priorizacion/>

*Servicios de informática en la nube | Microsoft Azure*. (2021). Obtenido de Azure:

<https://azure.microsoft.com/es-es/>

Shubhashree, M. (26 de Junio de 2018). *Core and Non-core Roles in Scrum*. Obtenido de

[https://www.linkedin.com/pulse/core-non-core-roles-scrum-m-](https://www.linkedin.com/pulse/core-non-core-roles-scrum-m-shubhashre)

[shubhashrehttps://www.linkedin.com/pulse/core-non-core-roles-scrum-m-shubhashre](https://www.linkedin.com/pulse/core-non-core-roles-scrum-m-shubhashre)

Sylvia Chou, W.-Y., Gaysynsky, A., & Cappella, J. (2021). Qué camino tomar: la información

errónea sobre la salud en las redes sociales. *Revista Panamericana de Salud Pública*.

UNESCO. (03 de 03 de 2021). *UNESCO construir la paz en la mente de los hombres y de las mujeres*. Obtenido de Moderación de contenidos y la protección de la libertad de expresión de

los usuarios de redes sociales: [https://es.unesco.org/news/moderacion-contenidos-y-](https://es.unesco.org/news/moderacion-contenidos-y-proteccion-libertad-expresion-usuarios-redes-sociales)

[proteccion-libertad-expresion-usuarios-redes-sociales](https://es.unesco.org/news/moderacion-contenidos-y-proteccion-libertad-expresion-usuarios-redes-sociales)



Vivanco, J. (05 de 06 de 2019). *El modelo C4 de documentación para la Arquitectura de Software*. Obtenido de <https://medium.com/@javiervivanco/el-modelo-c4-de-documentaci%C3%B3n-para-la-arquitectura-de-software-424704528390>

Wen-Ying, S., Gaysynsky, A., & Cappella, J. (2021). *Qué camino tomar: la información errónea sobre la salud en las redes sociales*.

Wilfrido Martel, D. C. (2018). *Tendencias de la computacion* . Obtenido de Determinación de niveles de agresividad en comentarios de la red social Facebook por medio de Minería de Texto: <https://journal.espe.edu.ec/ojs/index.php/geeks/article/view/280>

## Anexos