



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

## **DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN**

**CARRERA DE INGENIERÍA EN SISTEMAS E INFORMÁTICA**

**TRABAJO DE TITULACIÓN PREVIO LA OBTENCIÓN DEL  
TÍTULO DE INGENIERO EN SISTEMAS E INFORMÁTICA**

**TEMA: TUTORÍA ASISTIDA POR COMPUTADOR PARA LA  
REDACCIÓN DE ARTÍCULOS CIENTÍFICOS MEDIANTE EL  
PROCESAMIENTO DE LENGUAJE NATURAL**

**AUTORES: GARCÍA FLORES, JEAN PIERRE  
JIMÉNEZ PATIÑO, JOHANNA ESTEFANÍA**

**DIRECTOR: DR. FONSECA C., EFRAÍN R.**

**SANGOLQUÍ**

**2017**



**DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN  
CARRERA DE INGENIERÍA EN SISTEMAS E INFORMÁTICA**

**CERTIFICADO**

Certifico que el trabajo de titulación, "*TUTORIA ASISTIDA POR COMPUTADOR PARA LA REDACCIÓN DE ARTICULOS CIENTIFICOS MEDIANTE EL PROCESAMIENTO DE LENGUAJE NATURAL*" realizado por la señorita *JOHANNA ESTEFANIA JIMENEZ PATINO*, ha sido revisado en su totalidad y analizado por el software anti-plagio, el mismo cumple con los requisitos teóricos, científicos, técnicos, metodológicos y legales establecidos por la Universidad de las Fuerzas Armadas ESPE, por lo tanto me permito acreditarlo y autorizar a la señorita *JOHANNA ESTEFANIA JIMENEZ PATINO* para que lo sustente públicamente.

**Sangolquí, 24 de agosto del 2017**

---

DR. EFRAIN R. FONSECA C.  
**DIRECTOR**



**DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN  
CARRERA DE INGENIERÍA EN SISTEMAS E INFORMÁTICA**

**CERTIFICADO**

Certifico que el trabajo de titulación, "*TUTORÍA ASISTIDA POR COMPUTADOR PARA LA REDACCIÓN DE ARTÍCULOS CIENTÍFICOS MEDIANTE EL PROCESAMIENTO DE LENGUAJE NATURAL*" realizado por el señor **JEAN PIERRE GARCÍA FLORES**, ha sido revisado en su totalidad y analizado por el software anti-plagio, el mismo cumple con los requisitos teóricos, científicos, técnicos, metodológicos y legales establecidos por la Universidad de las Fuerzas Armadas ESPE, por lo tanto me permito acreditarlo y autorizar al señor **JEAN PIERRE GARCÍA FLORES** para que lo sustente públicamente.

**Sangolquí, 24 de agosto del 2017**

---

**DR. EFRAÍN R. FONSECA C.  
DIRECTOR**



**DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN  
CARRERA DE INGENIERÍA EN SISTEMAS E INFORMÁTICA**

**AUTORIA DE RESPONSABILIDAD**

Yo, **JOHANNA ESTEFANIA JIMENEZ PATINO**, con cédula de identidad N° 1721674016, declaro que este trabajo de titulación "**TUTORIA ASISTIDA POR COMPUTADOR PARA LA REDACCION DE ARTICULOS CIENTIFICOS MEDIANTE EL PROCESAMIENTO DE LENGUAJE NATURAL**" ha sido desarrollado considerando los métodos de investigación existentes, así como también se ha respetado los derechos intelectuales de terceros considerándose en las citas bibliográficas.

Consecuentemente declaro que este trabajo es de mi autoría, en virtud de ello me declaro responsable del contenido, veracidad y alcance de la investigación mencionada.

**Sangolquí, 24 de agosto de 2017**

---

JOHANNA ESTEFANIA JIMENEZ PATINO

C.C. 1721674016



**DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN  
CARRERA DE INGENIERÍA EN SISTEMAS E INFORMÁTICA**

**AUTORÍA DE RESPONSABILIDAD**

Yo, **JEAN PIERRE GARCÍA FLORES**, con cédula de identidad N° 1723625909, declaro que este trabajo de titulación "**TUTORÍA ASISTIDA POR COMPUTADOR PARA LA REDACCIÓN DE ARTÍCULOS CIENTÍFICOS MEDIANTE EL PROCESAMIENTO DE LENGUAJE NATURAL**" ha sido desarrollado considerando los métodos de investigación existentes, así como también se ha respetado los derechos intelectuales de terceros considerándose en las citas bibliográficas.

Consecuentemente declaro que este trabajo es de mi autoría, en virtud de ello me declaro responsable del contenido, veracidad y alcance de la investigación mencionada.

**Sangolquí, 24 de agosto de 2017**

---

**JEAN PIERRE GARCÍA FLORES**

C.C. 1723625909



**DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN  
CARRERA DE INGENIERÍA EN SISTEMAS E INFORMÁTICA**

**AUTORIZACION**

Yo, **JOHANNA ESTEFANIA JIMENEZ PATINO**, autorizo a la Universidad de las Fuerzas Armadas ESPE publicar en la biblioteca Virtual de la institución el presente trabajo de titulación "**TUTORIA ASISTIDA POR COMPUTADOR PARA LA REDACCION DE ARTICULOS CIENTIFICOS MEDIANTE EL PROCESAMIENTO DE LENGUAJE NATURAL**" cuyo contenido, ideas y criterios son de mi autoría y responsabilidad.

**Sangolquí, 24 de agosto de 2017**

---

JOHANNA ESTEFANIA JIMENEZ PATINO

C.C. 1721674016



**DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN  
CARRERA DE INGENIERÍA EN SISTEMAS E INFORMÁTICA**

**AUTORIZACIÓN**

Yo, **JEAN PIERRE GARCÍA FLORES**, autorizo a la Universidad de las Fuerzas Armadas ESPE publicar en la biblioteca Virtual de la institución el presente trabajo de titulación "**TUTORÍA ASISTIDA POR COMPUTADOR PARA LA REDACCIÓN DE ARTÍCULOS CIENTÍFICOS MEDIANTE EL PROCESAMIENTO DE LENGUAJE NATURAL**" cuyo contenido, ideas y criterios son de mi autoría y responsabilidad.

**Sangolquí, 24 de agosto de 2017**

A handwritten signature in blue ink, which appears to read 'JP García Flores', is written over a horizontal dashed line.

**JEAN PIERRE GARCÍA FLORES**

**C.C: 1723625909**

## DEDICATORIA

A Dios, quien supo guiar mi camino, darme fuerzas para salir adelante y no decaer frente a las adversidades que se presentaban en el transcurso de mi formación profesional.

A mi amada madre que es el pilar fundamental en mi vida, por su amor, sus consejos, por ser mi ejemplo de constancia y perseverancia, y por todo el esfuerzo, dedicación y cuidado con el que siempre estuvo pendiente de mí, dándome fuerzas y ánimos para no decaer. A mi padre y hermanos por estar presentes en los momentos difíciles y de alegría en mi vida.

A aquella persona que de manera incondicional siempre estuvo presente en los buenos y malos momentos, quien es una fuente de motivación, inspiración y felicidad.

Johanna Jiménez

## DEDICATORIA

Dedico este trabajo a mis padres y hermanos, quienes son la esencia de mí ser y que guiaron e influenciaron el camino que me ha permitido lograr la culminación de este estudio y con ello de mi carrera.

Jean Pierre García

## **AGRADECIMIENTO**

A Dios, por darme la vida y familia maravillosa que siempre me han apoyado en cada meta que me he propuesto.

A mi amada madre, por su ejemplo de lucha constante, por su infinito amor, por siempre apoyarme y hacer de mí una persona de bien, por su dedicación y paciencia. A mi padre y hermanos por confiar y creer siempre en mí.

A mi director de tesis, Dr. Efraín R. Fonseca, por brindarme su conocimiento, por su paciencia y motivación que han sido elementos vitales para la culminación de este trabajo y para mi formación como profesional e investigadora.

A todas esas personas que, de una u otra forma, colaboraron y acompañaron en el camino hacia mi meta profesional, quienes creyeron en mí y dieron paso al cumplimiento de este reto.

Johanna Jiménez

## **AGRADECIMIENTO**

Agradezco a mis profesores, especialmente a mi director de tesis, quienes formaron gran parte de mi criterio profesional y compartieron conmigo gratos momentos. Agradezco también a mi familia quienes siempre estuvieron pendientes de mi bienestar compartiendo mis alegrías y apoyándome en mis momentos difíciles. Finalmente, agradezco a mis amigos que estuvieron cuando lo necesité, a quienes pude enseñar y de quienes pude aprender grandes lecciones.

Jean Pierre García

**ÍNDICE**

<b>CERTIFICADO .....</b>	<b>ii</b>
<b>CERTIFICADO .....</b>	<b>iii</b>
<b>AUTORÍA DE RESPONSABILIDAD .....</b>	<b>iv</b>
<b>AUTORÍA DE RESPONSABILIDAD .....</b>	<b>v</b>
<b>AUTORIZACIÓN .....</b>	<b>vi</b>
<b>AUTORIZACIÓN .....</b>	<b>vii</b>
<b>DEDICATORIA .....</b>	<b>viii</b>
<b>DEDICATORIA .....</b>	<b>ix</b>
<b>AGRADECIMIENTO .....</b>	<b>x</b>
<b>AGRADECIMIENTO .....</b>	<b>xi</b>
<b>ÍNDICE .....</b>	<b>xii</b>
<b>ÍNDICE DE TABLAS .....</b>	<b>xv</b>
<b>ÍNDICE DE FIGURAS .....</b>	<b>xvi</b>
<b>RESUMEN .....</b>	<b>xvii</b>
<b>ABSTRACT .....</b>	<b>xviii</b>
<b>CAPÍTULO I .....</b>	<b>1</b>
<b>INTRODUCCIÓN .....</b>	<b>1</b>
1.1. Antecedentes .....	1
1.2. Problemática .....	3
1.3. Justificación .....	5
1.4. Objetivos .....	6
1.4.1. Objetivo General .....	6
1.4.2. Objetivos Específicos .....	6

1.5. Alcance .....	7
<b>CAPÍTULO II.....</b>	<b>10</b>
<b>ESTADO DEL ARTE .....</b>	<b>10</b>
2.1. Método.....	10
2.1.1. Definición de los criterios de inclusión y exclusión .....	10
2.1.2. Selección de literatura base .....	11
2.1.3. Validación cruzada .....	12
2.1.4. Búsqueda de estudios candidatos .....	14
2.1.5. Depuración de estudios.....	15
2.2. Descripción de estudios del estado del arte .....	16
<b>CAPÍTULO III.....</b>	<b>27</b>
<b>METODOLOGÍA Y MARCO TEÓRICO.....</b>	<b>27</b>
3.1. Metodología .....	27
3.1.1. Design science .....	27
3.1.2. Amenazas a la validez .....	31
3.2. Marco teórico .....	31
3.2.1. Fundamentación Filosófica .....	31
3.2.2. Hipótesis .....	33
3.2.3. Red de Categorías .....	33
3.2.4. Fundamentación Científica De La Variable Independiente .....	34
3.2.5. Fundamentación Científica de la Variable Dependiente.....	42
<b>CAPÍTULO IV .....</b>	<b>44</b>
<b>DISEÑO Y DESARROLLO.....</b>	<b>44</b>
4.1. DISEÑO .....	44

4.1.1. Diseño del modelo.....	44
4.2. DESARROLLO .....	53
4.2.1. Descripción de la técnica .....	53
<b>CAPÍTULO V .....</b>	<b>57</b>
<b>VALIDACIÓN DEL MODELO .....</b>	<b>57</b>
5.1. Pruebas .....	58
<b>CAPÍTULO VI .....</b>	<b>63</b>
<b>CONCLUSIONES Y FUTURAS LÍNEAS DE INVESTIGACIÓN .....</b>	<b>63</b>
6.1. Conclusiones .....	63
6.2. Futuras Líneas de Investigación .....	64
<b>REFERENCIAS BIBLIOGRÁFICAS .....</b>	<b>66</b>

**ÍNDICE DE TABLAS**

Tabla 1 Artículos seleccionados por los investigadores.....	12
Tabla 2 Artículos resultantes del proceso de validación cruzada.....	13
Tabla 3 Artículos científicos obtenidos de la base digital Scopus .....	15
Tabla 4 Frecuencia de movimientos observados en el resumen de cada artículo .....	48
Tabla 5 Número de caracteres encontrado.....	52

## ÍNDICE DE FIGURAS

Figura 1 Fases para el desarrollo de la investigación .....	7
Figura 2 Arquitectura de la GLN .....	8
Figura 3 Arquitectura de los STI .....	9
Figura 4 Proceso de Estrategia de Búsqueda de Artículos Científicos .....	10
Figura 5 Ciclo de ingeniería .....	28
Figura 6 Red de Categorías de las Variables de Investigación .....	33
Figura 7 Interacción de los Módulos de un STI.....	36
Figura 8 Arquitectura de un Sistema de GLN .....	45
Figura 9 Mensajes a transmitir en su respectivo orden.....	46
Figura 10 Combinación de arquitecturas .....	47
Figura 11 Orden encontrado según la estructura de los artículos.....	52
Figura 12 Interfaz del prototipo .....	53
Figura 13 Preguntas generadas.....	54
Figura 14 Panel de identificación de secciones .....	55
Figura 15 Mensaje de número límite de caracteres de escritura .....	56
Figura 16 Texto resultante del proceso de escritura .....	56
Figura 17 Test de validación. Sujeto de pruebas 1 .....	58
Figura 18 Test de validación. Sujeto de pruebas 2 .....	59
Figura 19 Test de validación. Sujeto de pruebas 3.....	61
Figura 20 Test de validación. Sujeto de pruebas 4 .....	62

## RESUMEN

La investigación científica permite ampliar el conocimiento actual de los seres humanos mediante el hallazgo de nuevos descubrimientos utilizando, por ejemplo, la publicación de artículos científicos. Las revistas científicas han desarrollado un estilo, una estructura y una organización distintivos para constatar la calidad de un artículo. Es por esto que para abordar el complejo proceso de escritura científica se han desarrollado nuevos enfoques de enseñanza interdisciplinaria utilizando la Generación de Lenguaje Natural (GLN). Además de la GLN existen los Sistemas de Tutoría Inteligente que con el uso de preguntas, son una forma potencial de apoyo para los sistemas de aprendizaje de escritura. Por lo tanto, este estudio tiene como propósito construir un modelo de tutoría asistida por computador para la redacción de artículos científicos, mediante el PLN. Se realizó una revisión de literatura que proporcionó información sobre el proceso de escritura científica, luego, se efectuaron preguntas a investigadores experimentados para obtener la bases que permiten continuar con un análisis a un corpus de artículos científicos escogidos para extraer patrones acordes a la GLN, con lo cual, se generó un modelo lingüístico que sirvió para la creación de un diálogo basado en los Sistemas de Tutoría Inteligente. Como resultado se diseñó una arquitectura que integra la GLN y los Sistemas de Tutoría Inteligente que responde al dominio de la redacción de artículos científicos. Finalmente, se realizó un prototipo que refleja los resultados obtenidos en el estudio.

### **PALABRAS CLAVE:**

- **ESCRITURA CIENTÍFICA**
- **PROCESAMIENTO DEL LENGUAJE NATURAL**
- **GENERACIÓN DEL LENGUAJE NATURAL**
- **SISTEMAS DE TUTORÍA ASISTIDA**
- **PUBLICACIÓN DE ARTÍCULOS**

## ABSTRACT

Scientific research allows us to extend the current knowledge of human beings by finding new discoveries using, for example, the publication of scientific articles. Scientific journals have developed a distinctive style, structure and organization to verify the quality of an article. That is why new approaches to interdisciplinary teaching using the Natural Language Generation (GLN) have been developed to address the complex process of scientific writing. In addition to GLN, there are Intelligent Tutoring Systems that, with the use of questions, are a potential form of support for writing learning systems. Therefore, this study aims to build a model of computer-assisted tutoring for the writing of scientific articles, using the PLN. A literature review was conducted that provided information about the scientific writing process, then questions were asked to experienced researchers to obtain the basis for an analysis of a corpus of scientific papers selected to extract GLN-compliant patterns, with Which generated a linguistic model that served to create a dialogue based on Intelligent Tutoring Systems. As a result, an architecture was designed that integrates the GLN and the Intelligent Tutoring Systems that responds to the domain of the writing of scientific articles. Finally, a prototype was performed that reflects the results obtained in the study.

### KEYWORDS:

- **SCIENTIFIC WRITING**
- **NATURAL LANGUAGE PROCESSING**
- **NATURAL LANGUAGE GENERATION**
- **ASSISTED TUTORING SYSTEMS**
- **PUBLICATION OF ARTICLES**

## CAPÍTULO I

### INTRODUCCIÓN

#### 1.1. Antecedentes

La investigación científica se forja en las universidades. Las facultades de ingeniería, por ejemplo, contribuyen significativamente al desarrollo de la tecnología en todos los campos (Jayalakshmi, Suvarna, Vidya, & Sujatha, 2015), lo cual es posible a través de la investigación. La investigación científica es un proceso metódico que permite ampliar el conocimiento actual de los seres humanos, la tecnología y la ciencia, la cultura y la sociedad mediante el hallazgo de nuevos descubrimientos (Jayalakshmi, Suvarna, Vidya, & Sujatha, 2015).

La principal forma de difundir los hallazgos científicos es a través de reportes de investigación, dado que una investigación no se da a conocer hasta que sus resultados sean publicados (Mantilla-Villareal, y otros, 2010). Aunque no es la única manera, el artículo científico es la principal forma de comunicación en la comunidad científica. Los artículos pasan generalmente por un proceso de revisión por pares, en el cual se constata su calidad y se verifica que contengan los elementos necesarios para fortalecer el conocimiento científico en una temática particular (Leyva, 2013).

Para constatar la calidad de un artículo, en los últimos 350 años las revistas científicas han desarrollado un estilo, una estructura y una organización distintivos, para que el investigador pueda centrarse en lo que mejor sabe, la ciencia, y preocuparse menos por la escritura. Estas restricciones permiten al autor y al lector centrarse en el contenido, lo que a menudo resulta en un mejor artículo (Mack, 2014).

La escritura efectiva es vital para mejorar la calidad de un artículo y una habilidad crítica para los estudiantes universitarios (Geithner & Pollastro, 2015). Hay varios criterios a tomar en cuenta para que la escritura sea efectiva; por ejemplo, escribir de forma clara, breve y evitando redundancias para que el lector

entienda mejor y más rápido; emplear correctamente los verbos y utilizar un léxico variado (Orwell's, 2007). Existen varios estudios y libros que abordan el proceso de redacción de artículos científicos; sin embargo, se han reportado pocos esfuerzos para desarrollar de forma sistemática las habilidades de redacción en el investigador (Cargill & Smernik, 2015).

Un enfoque de enseñanza interdisciplinaria para abordar el complejo proceso de escritura científica se ha desarrollado recientemente utilizando la investigación en lingüística aplicada, con resultados exitosos (Cargill & Smernik, 2015). Dentro de la Lingüística Aplicada se encuentra la Lingüística Computacional también llamada Procesamiento del Lenguaje Natural (PLN) (Vicente, Barros, Peregrino, Agulló, & Lloret, 2015). El PLN juega un papel importante en cómo las computadoras interpretan y procesan los lenguajes naturales humanos a fin de desarrollar aplicaciones útiles de aprendizaje (Al Emran & Shaalan, 2014).

La Generación de Lenguaje Natural (GLN) es una disciplina del PLN, cuyo objetivo principal es investigar acerca de la creación de aplicaciones informáticas que tengan la capacidad de producir textos de alta calidad en lenguaje natural (Vicente, Barros, Peregrino, Agulló, & Lloret, 2015). Durante las últimas décadas se ha generado mucha investigación con relación a los correctores ortográficos y gramaticales, para dar soporte a las aplicaciones de GLN, cuyos resultados se han integrado en varios programas de procesamiento de texto. En la práctica, estas herramientas podrían alertar a los usuarios en sus errores de escritura, pero no podrían ayudar a los usuarios a escribir un artículo desde cero (Liu, Lee, Yu, & Chen, 2010).

Se han propuesto muchos sistemas para apoyar el proceso de escritura. Por ejemplo: Roscoe & otros (2014) proponen el sistema W-Pal, el cual es un sistema diseñado para mejorar la competencia de escritura de los estudiantes a través de una combinación basada en juegos, redacción de ensayos y retroalimentación automatizada.

El uso de preguntas es una forma potencial de apoyo para los sistemas de aprendizaje de escritura (Liu, Calvo, & Rus, 2012). En el estudio realizado por

Liu, Calvo, Aditomo y Pizzato (2012) se muestran algunas de las preguntas genéricas para apoyar la escritura científica. Además, proponen un sistema automático para la generación de preguntas. En dicho estudio se evalúa la utilidad de las preguntas generadas por su algoritmo y se las compara con las preguntas genéricas.

Como se puede apreciar, la escritura científica es un proceso complejo y de continuo estudio (Mantilla-Villareal, y otros, 2010), (Leyva, 2013), (Geithner, 2016), (Cargill & Smernik, 2015), (Jayalakshmi, Suvarna, Vidya, & Sujatha, 2015), (Mack, 2016). Las herramientas informáticas podrían facilitar la comprensión y la ejecución de muchos procesos, por ejemplo, la escritura automática es un tema abordado en la disciplina de la GLN, donde sus procesos están bastante organizados (Al Emran & Shaalan, 2014), (Vicente, Barros, Peregrino, Agulló, & Lloret, 2015). Las preguntas de activación son formas potenciales de apoyo para aprender a escribir (Liu, Calvo, & Rus, 2012). En este documento se plantea utilizar las bondades del cuestionamiento automático para apoyar a la escritura científica, basándose en los procesos de la GNL, las cuales proporcionan al mismo tiempo el sustento para construir un sistema de escritura asistida por computadora.

## **1.2. Problemática**

Esencialmente un manuscrito científico se mantiene unido a un área de investigación, es decir que si se sigue un enfoque riguroso en cada etapa de la investigación desde la hipótesis inicial hasta el diseño, la ejecución y el análisis de los experimentos, el estudio se enfrentará en última instancia al escrutinio de los críticos más duros y de la comunidad científica. Pero por el contrario si un manuscrito es incoherente puede socavar incluso el estudio científico más meticuloso e impedir la comunicación de un buen estudio científico (Gemayel, 2016).

Por ejemplo, en el estudio de López Leyva de 2013 se indica que:

La preparación del manuscrito debe atenderse en cuatro apartados. En primer lugar, es necesario definir un plan de escritura, considerando el tiempo y los recursos disponibles; después definir la etapa inicial de preparación, es decir, cómo iniciar el escrito; atender los procedimientos de edición y, por último, llevar a cabo la revisión final (pág. 15).

En el Ecuador la escritura de artículos científicos es un proceso relativamente nuevo, por esta razón los estudiantes no tienen experiencia en el proceso de escritura científica ni en su planificación (Loyola Illescas, y otros, 2014). El apoyo de investigadores experimentados es fundamental para una mejor comprensión de la planificación del artículo. El problema del apoyo de investigadores experimentados radica en su costo (Loyola Illescas, y otros, 2014).

La escritura es inherentemente un proceso creativo. Eso parecería apropiado para un investigador, donde la creatividad junto con pensamiento crítico es la llave al éxito. Por desgracia, muchos científicos consideran la tarea de escritura como intimidante (Mack, 2014).

El proceso de comunicación se realiza a través del lenguaje. Para conseguirlo, se utilizan tareas de alto nivel cognitivo cuya complejidad se pone de manifiesto en la tarea de automatizar el proceso (Vicente, Barros, Peregrino, Agulló, & Lloret, 2015). Aunque la escritura es muy importante, escribir un artículo desde cero es un trabajo difícil para muchos (Liu, Lee, Yu, & Chen, 2010). La escritura es un proceso complejo que implica varias etapas tales como: planificación, edición y revisión. Este proceso podría ser apoyado por herramientas computacionales que evalúen la escritura, proporcionando instrucciones a los estudiantes para mejorar sus habilidades de escritura (Gorrostieta & López-López, 2016).

Una vez que se ha descrito el problema abordado en esta investigación, es necesario analizar las propuestas existentes sobre esta temática para solucionarlo. Se ha descrito el problema de la dificultad en el proceso de escritura y por ende el costo que implica contar con el tiempo de tutores experimentados,

por lo cual se focalizará la búsqueda de soluciones tecnológicas que aborden estos problemas. En secciones posteriores se intenta resumir los estudios que proponen mejorar el proceso de escritura, de la mano de la organización y generación de ideas, utilizando para esto la tecnología.

### **1.3. Justificación**

Una de las principales tareas de las universidades es el proceso de difusión de sus contribuciones científicas generadas en proyectos de investigación. La publicación de artículos es la forma usual, a través de la cual, las universidades exponen sus resultados de investigación; por lo tanto, es fundamental que los investigadores tengan la suficiente habilidad para que los resultados de investigación sean presentados de una manera adecuada y concisa. Una forma de generar las características idóneas para que los investigadores dominen la escritura de artículos es creando cursos y plataformas que asesoren y brinden una guía tecnológica que construya de forma asistida artículos científicos para los investigadores. No obstante, a pesar de la gran ayuda que representa brindar éstas facilidades, el costo que involucra capacitar a todos los estudiantes y profesores en la escritura científica es sumamente elevado, aunque necesario para mejorar las producciones científicas, por lo tanto, una alternativa para abordar este problema es utilizar tutores asistidos por computadora.

En el presente proyecto se pretende sentar un precedente para el uso de estos tutores en el apoyo de la escritura de artículos, intentando fomentar la creación de herramientas de este tipo que soporten la escritura científica. Con la creación de tutores inteligentes para la escritura, las universidades pueden ahorrar muchos recursos y aumentar su eficiencia en la publicación de artículos.

Se ha demostrado que el cuestionamiento mejora los resultados del aprendizaje, y la generación automática de preguntas puede facilitar en gran medida esta tarea, como los sistemas de tutoría inteligente.

La mayoría de los Sistemas de Preguntas Automáticas como los Sistemas de Tutoría Inteligente, usan software de análisis y algoritmos de transformación para crear preguntas; utilizando el entendimiento del lenguaje natural en el proceso de generación de lenguaje natural (GLN) se puede crear plantillas de preguntas que se emparejen a lo que la oración está comunicando para generar preguntas de mayor calidad (Mazidi, 2016), se puede mostrar la utilidad de este nuevo enfoque, implementando un sistema específico para un proceso complejo como es la escritura de artículos científicos en lenguaje natural.

En este proyecto se presenta un prototipo de herramienta que utiliza los métodos de los Sistemas de Tutoría Inteligente y los procesos de la Generación del Lenguaje Natural (PLN), para la creación de preguntas que guíen el proceso de escritura. Esta herramienta pretende fomentar la cooperación entre el investigador y el computador en el proceso de escritura.

#### **1.4. Objetivos**

##### **1.4.1. Objetivo General**

Construir un modelo de tutoría asistida por computador para la redacción de artículos científicos, mediante el Procesamiento de Lenguaje Natural.

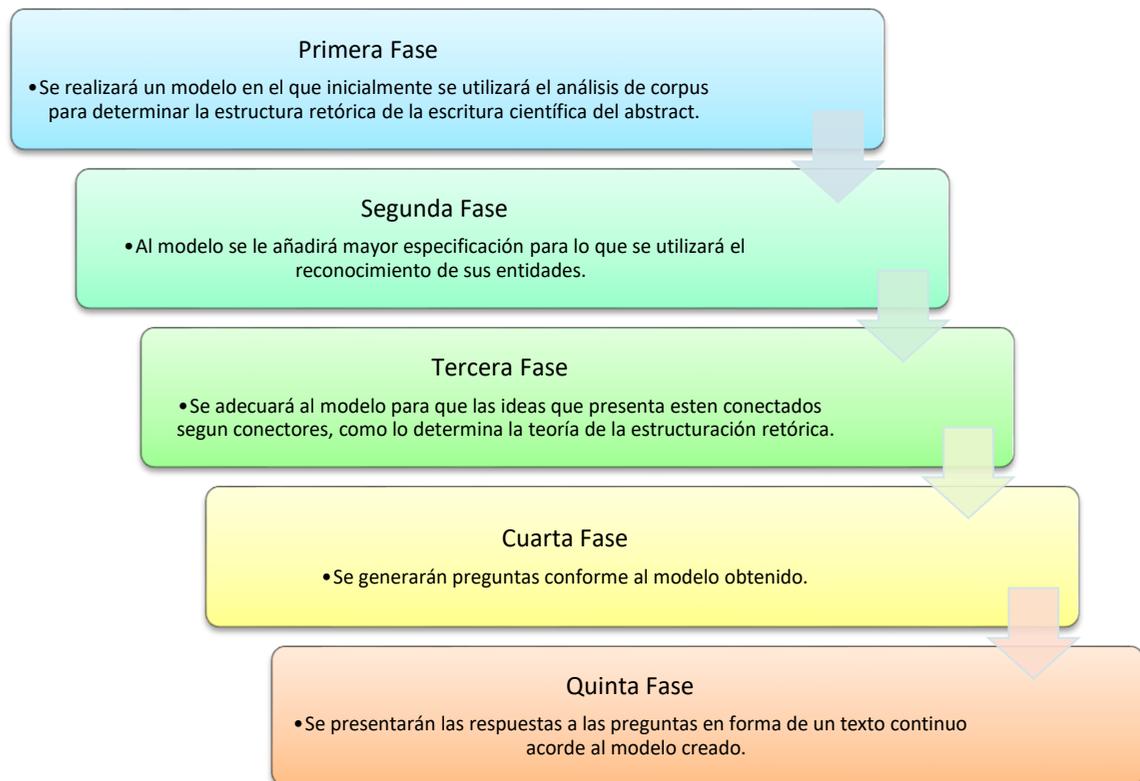
##### **1.4.2. Objetivos Específicos**

- Identificar patrones en la estructuración de la escritura retórica de los artículos científicos para formalizar un modelo organizado de expresiones.
- Determinar las entidades utilizadas en las secciones identificadas en la estructuración de la escritura retórica de los artículos científicos.
- Identificar patrones de términos que conectan las secciones identificadas en la estructuración de los artículos según la teoría de la estructuración retórica.

- Diseñar un árbol de preguntas relacionadas con los patrones lingüísticos identificados.
- Estructurar las respuestas obtenidas del árbol de preguntas para la formulación de sugerencias en la realización del contenido.
- Construir un prototipo que facilite la escritura del resumen del artículo científico.
- Validar el prototipo con investigadores del Grupo de Investigación en Modelos de Producción de Software e Industria Inteligente (GrIMPESOFTII).

### 1.5. Alcance

Para cumplir con los objetivos planteados, se abordará cada proceso en cinco fases distintas (ver figura 1).

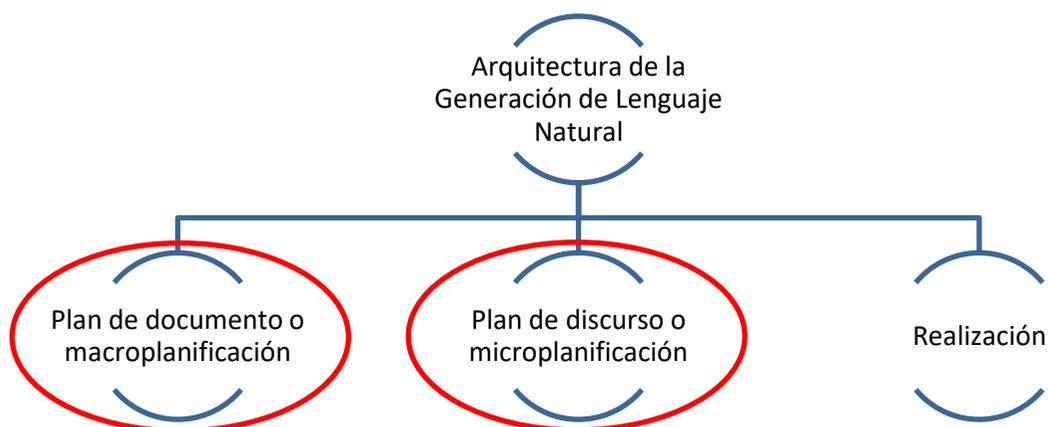


**Figura 1 Fases para el desarrollo de la investigación**

Adicionalmente, se han planteado preguntas de investigación que delimitan y guían el desarrollo de la presente tesis:

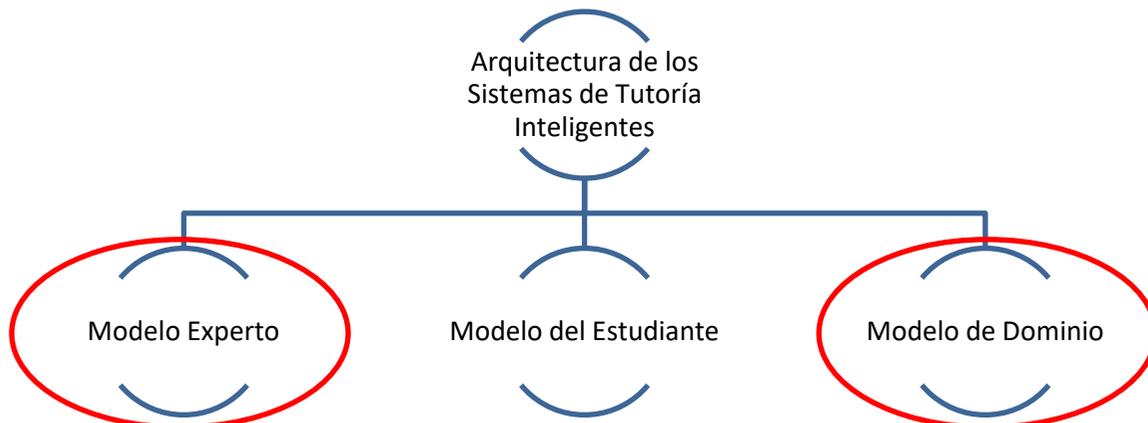
- ¿El pensamiento informático o la abstracción del mundo real puede apoyar el entendimiento del proceso de escritura científica?
- ¿Un diálogo computarizado puede facilitar la escritura científica?
- ¿Dividir el proceso de escritura científica en las fases que determina la GLN facilita el proceso de escritura científica?
- ¿Puede asignarse roles tanto a humanos como a computadoras para trabajar en conjunto en el proceso de escritura científica?

A continuación, se describen dos arquitecturas que serán tomadas de referencia para el desarrollo de la propuesta planteada. La primera es la arquitectura de los procesos de la Generación del Lenguaje Natural (ver figura 2).



**Figura 2 Arquitectura de la GLN**

Para el presente estudio, se desarrollarán los dos primeros módulos: la macroplanificación y la microplanificación, los cuales serán descritos posteriormente en el marco teórico.



**Figura 3 Arquitectura de los STI**

La segunda arquitectura que se tomará de referencia es la de los Sistemas de Tutoría Inteligente (ver figura 3), los módulos que se realizarán de esta arquitectura son: Modulo del Experto y Modelo del Dominio.

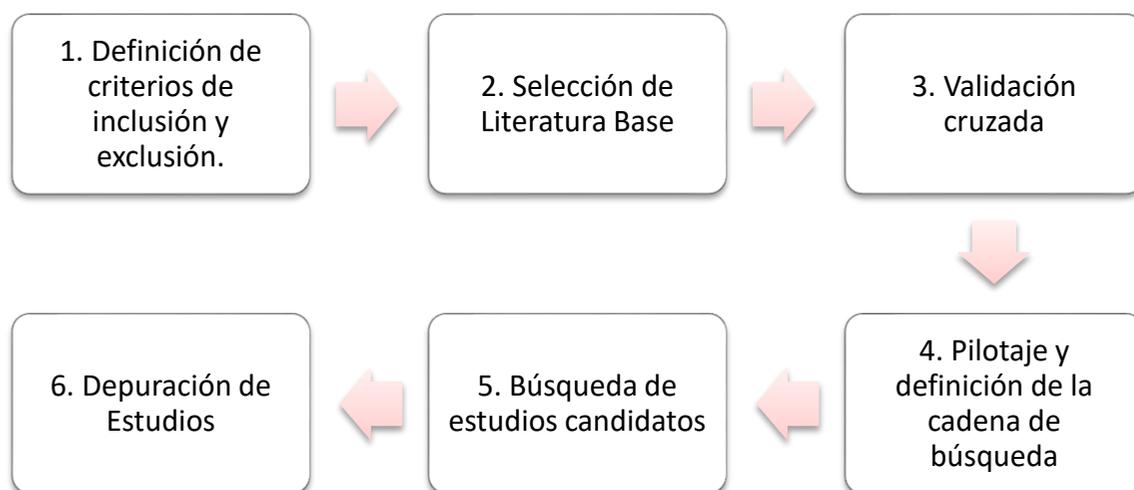
Las dos arquitecturas interactuarán entre sí, de tal manera forma que las respuestas al diálogo del Sistema Inteligente de Tutoría servirán como entrada para el Generador de Lenguaje Natural.

## CAPÍTULO II

### ESTADO DEL ARTE

#### 2.1. Método

Para la búsqueda de estudios circunscritos en la problemática de interés se siguió un método ad-hoc, obteniéndose como resultado 23 artículos, los cuales constituyeron la base para el estudio del estado del arte. La figura 4 muestra las etapas del método de búsqueda efectuado, las cuales se describen a detalle más adelante.



**Figura 4 Proceso de Estrategia de Búsqueda de Artículos Científicos**

##### 2.1.1. Definición de los criterios de inclusión y exclusión

Los criterios de inclusión y exclusión constituyen los parámetros que deben cumplir los estudios para ser considerados o no en la investigación; estos criterios ayudan a ubicar la información pertinente.

**Criterios de Inclusión:** Los criterios de inclusión definidos para la presente estrategia de búsqueda son los siguientes:

- El artículo describe información sobre Escritura Científica.
- Artículos que contengan propuestas sobre Escritura Asistida por Computador.
- Artículos que propongan herramientas para dar solución a los problemas de escritura.
- El artículo describe los problemas suscitados al momento de escribir un artículo científico.
- Artículos científicos que estén publicados en revistas indexadas de alto impacto.

**Criterios de Exclusión:** Los criterios de exclusión definidos para la presente estrategia de búsqueda son los siguientes:

- Artículos en los que se proponga un modelo o sistema como solución a los problemas de escritura y cuya validez no sea evaluada.
- Artículos que estén escritos en otro idioma que no sea inglés y que hayan sido publicados antes del 2012.
- Artículos que no respondan a ninguno de los criterios seleccionados para el presente estudio.

### **2.1.2. Selección de literatura base**

La etapa de selección de literatura base corresponde a la selección de estudios que responden a las características y objetivos planteados en la presente investigación. Además, el propósito de esta etapa es identificar palabras clave para posteriormente construir la cadena de búsqueda. Los estudios escogidos en esta sección no necesariamente conformarán el grupo de literatura base, ya que para ser considerados como tal deben atravesar un proceso de selección riguroso que será detallado más adelante.

La selección de la literatura base corresponde a una búsqueda individual por cada uno de los investigadores participantes, en este caso dos. Cada investigador seleccionó 6 artículos, dando un total de 12 artículos seleccionados

(ver Tabla 1), los cuales posteriormente fueron sometidos a un proceso de validación cruzada entre los mismos investigadores.

**Tabla 1**

**Artículos seleccionados por los investigadores**

AC	TÍTULO
<b>AC1</b>	The Sentence Fairy: a natural –language system to support children´s essay writing
<b>AC2</b>	Discourse Planning with an N-gram Model of Relations
<b>AC3</b>	G-Asks: An Intelligent Automatic Questions Generation System for Academic Writing Support
<b>AC4</b>	Automatic Question Generation: From NLU to NLG
<b>AC5</b>	Using Wikipedia and Conceptual Graph Structures to Generate Questions for Academic Writing Support
<b>AC6</b>	Intelligent Sentence Writing Tutor: A System Development Cycle
<b>AC7</b>	Automatic Question Generation system
<b>AC8</b>	Automatic question generation for supporting argumentation
<b>AC9</b>	Intelligent Tutoring with Natural Language Support in The BEETLE II System
<b>AC10</b>	An Intelligent Automatic Question Generation System for Academic Writing Support
<b>AC11</b>	Enhancing Writing through Strengthened Executive Function
<b>AC12</b>	Learning to Order Facts for Discourse Planning in Natural Language Generation

**2.1.3. Validación cruzada**

En esta etapa, los artículos encontrados por cada investigador fueron sometidos a un proceso de validación cruzada entre los mismos investigadores. De esta forma se garantiza que los artículos seleccionados cumplan los criterios

de inclusión y exclusión y sean consensuados por los investigadores. La validación cruzada consiste en hacer un grupo focal para dirimir acerca de si los artículos propuestos por cada uno de los investigadores sirven como base para la construcción de la cadena de búsqueda. Los artículos resultantes de este proceso fueron nueve (ver Tabla 2).

**Tabla 2**

**Artículos resultantes del proceso de validación cruzada**

AC	TÍTULO
<b>AC1</b>	The Sentence Fairy: a natural –language system to support children´s essay writing
<b>AC2</b>	Discourse Planning with an N-gram Model of Relations
<b>AC3</b>	G-Asks: An Intelligent Automatic Questions Generation System for Academic Writing Support
<b>AC4</b>	Automatic Question Generation: From NLU to NLG
<b>AC5</b>	Using Wikipedia and Conceptual Graph Structures to Generate Questions for Academic Writing Support
<b>AC6</b>	<del>Intelligent Sentence Writing Tutor: A System Development Cycle</del>
<b>AC7</b>	Automatic Question Generation system
<b>AC8</b>	Automatic question generation for supporting argumentation
<b>AC9</b>	Intelligent Tutoring with Natural Language Support in The BEETLE II System
<b>AC10</b>	<del>An Intelligent Automatic Question Generation System for Academic Writing Support</del>
<b>AC11</b>	Enhancing Writing through Strengthened Executive Function
<b>AC12</b>	<del>Learning to Order Facts for Discourse Planning in Natural Language Generation</del>

#### 2.1.4. Búsqueda de estudios candidatos

De los artículos que resultaron de la validación cruzada se ubicaron términos que están relacionados con el propósito de la presente investigación y que sirven para la construcción de la cadena de búsqueda; dichos términos se indican a continuación:

- Computer Assisted Writing
- Writing Support
- Computer aided instruction
- Intelligent Call
- Natural Language Processing
- Natural Language Generation
- Intelligent Tutoring System

Con los términos encontrados en los artículos resultantes del proceso de validación cruzada, se construyeron varias cadenas de búsqueda con el propósito de obtener artículos relacionados al tema de estudio. Cada cadena fue piloteada en la base digital Scopus para validar su efectividad. La primera cadena de búsqueda propuesta fue:

(( ( computer assisted writing ) OR ( writing support ) OR ( intelligent call ) ) AND ( natural language generation ) ).

Esta cadena expuso varios artículos que no cumplían a cabalidad con el objeto de estudio, por lo cual se continuó realizando varios intentos que finalmente dieron paso a una cadena de búsqueda que permitió encontrar artículos científicos que cumplan con los criterios de inclusión y exclusión. La cadena utilizada fue la siguiente:

(( ( computer assisted writing ) OR ( writing support ) OR ( computer aided instruction ) OR ( intelligent call ) OR ( Intelligent Tutoring Systems ) ) AND ( ( natural language processing ) OR ( natural language generation ) ) ).

La cadena de búsqueda dio como resultado 267 artículos.

### 2.1.5. Depuración de estudios

Al total de artículos obtenidos, se le aplicó un primer filtro estableciendo la búsqueda de artículos publicados en los últimos 5 años, obteniéndose 127 artículos. Luego, se aplicó un filtro para obtener artículos que hayan sido publicados en revistas indexadas de alto impacto, por ejemplo: Expert Systems with Applications, Natural Language Engineering, Computational Linguistics, Computers in Human Behavior, Web Semantics. Posteriormente, los investigadores estudiaron cada uno de los artículos para obtener los artículos más afines al problema identificado, con lo cual se obtuvieron 18 estudios.

Por último, se añadieron artículos usando el proceso de referencia cruzada con lo que se obtuvieron 23 artículos (ver tabla 3) relacionados al tema de investigación y que cumplen con los requisitos de inclusión y exclusión.

**Tabla 3**

#### Artículos científicos obtenidos de la base digital Scopus

Artículo	Título
<b>AC1</b>	The Sentence Fairy: a natural-language generation system to support children`s essay writing
<b>AC2</b>	Discourse Planning with an N-gram Model of Relations
<b>AC3</b>	G-Asks: An Intelligent Automatic Question Generation System for Academic Writing Support
<b>AC4</b>	Automatic Question Generation: From NLU to NLG
<b>AC5</b>	Using Wikipedia and Conceptual Graph Structures to Generate Questions for Academic Writing Support
<b>AC6</b>	Automatic Question Generation system
<b>AC7</b>	"Automatic question generation for supporting argumentation
<b>AC8</b>	Intelligent Tutoring with Natural Language Support in the Beetle II System
<b>AC9</b>	Enhancing Writing through Strengthened Executive Function



<b>AC10</b>	Linking the Thoughts Analysis of Argumentation Structures in Scientific Publications
<b>AC11</b>	Argumentation Identification for Academic Support in Undergraduate Writings
<b>AC12</b>	Intelligent computer assisted blog writing system
<b>AC13</b>	Towards Computer-Assisted Language Learning with Robots, Wikipedia and CogInfoCom
<b>AC14</b>	Automatic Generation and Ranking of Questions for Critical Review
<b>AC15</b>	An Intelligent System Framework for an Automated Language Tutoring Tool
<b>AC16</b>	Toward automatic inference of causal structure in student essays
<b>AC17</b>	Automatic Coding of Questioning Patterns in Knowledge-building Discourse
<b>AC18</b>	A collaborative planning model of intentional structure. Computational Linguistics
<b>AC19</b>	Textual complexity and discourse structure in Computer-Supported Collaborative Learning.
<b>AC20</b>	Identifying off-topic student essays without topic-specific training data
<b>AC21</b>	A case based reasoning model for multilingual language generation in dialogues
<b>AC22</b>	Computer assisted writing system
<b>AC23</b>	Argumentation mining

## 2.2. Descripción de estudios del estado del arte

### **AC1 (Harbusch, Itsova, Koch, & Kühner, 2008): The Sentence Fairy: a natural –language system to support children´s essay writing**

En este estudio, los autores desarrollaron un sistema de PLN que implementa una “conferencia de escritura virtual” para la enseñanza de alemán para niños de primaria. Aquí, se sigue un curso diferente al implementar la

tecnología de generación de lenguaje natural para evaluar y mejorar la calidad gramatical de la producción estudiantil. Basándose en una representación abstracta de la historia en construcción, todas las paráfrasis se generan de forma totalmente automática. El sistema genera ejercicios que permiten a los alumnos mejorar sus oraciones, sin embargo, no se aprovecha las respuestas de los estudiantes para crear la estructura de todo un párrafo.

### **AC2 (Biran & McKeown, 2015): Discourse Planning with an N-gram Model of Relations**

Las transiciones entre las relaciones del discurso son importantes para la coherencia, y esto puede ayudar en la generación del lenguaje; Los autores de este artículo proponen una solución para la planificación del discurso dentro de los sistemas de generación de concepto a texto y así determinar el orden de los mensajes y las relaciones de discurso entre ellos. Para mejorar este tipo de sistemas se puede crear preguntas debidamente organizadas para estructurar el orden de los mensajes que se desean transmitir.

### **AC3 (Liu, Calvo, & Rus, 2012): G-Asks: An Intelligent Automatic Question Generation System for Academic Writing Support**

En este estudio los autores proponen un sistema de retroalimentación electrónica para apoyar la escritura. Este artículo presenta un nuevo sistema automático de generación de preguntas, que genera preguntas específicas para apoyar el aprendizaje de los estudiantes a través de la escritura. Los autores identifican los tipos de preguntas más frecuentes realizadas por los supervisores humanos. En este artículo no se genera una directiva para organizar las preguntas generadas.

### **AC4 (Mazidi & Tarau, 2016): Automatic Question Generation: From NLU to NLG**

El enfoque descrito en este estudio utiliza la comprensión del lenguaje natural (NLU) en el proceso de generación de lenguaje natural (NLG) analizando

primero el contenido semántico central de cada cláusula independiente en cada oración. Entonces las plantillas de la pregunta se emparejan a lo que la oración está comunicando para generar preguntas de mayor calidad. Este enfoque generó un mayor porcentaje de preguntas aceptables que los sistemas de generación de preguntas que usan un software de análisis y algoritmos de transformación para crear preguntas. Este enfoque puede ser utilizado para luego ser revisado por un experto y crear una mejor organización de ideas, para lo cual se puede aprovechar las respuestas a las preguntas generadas. Otro enfoque encontrado es la construcción de conocimiento a través de entornos colaborativos.

**AC5 (Liu M. , Calvo, Aditomo, & Pizzato, 2012): Using Wikipedia and Conceptual Graph Structures to Generate Questions for Academic Writing Support**

Los autores presentan un enfoque para la generación de preguntas semiautomáticas para apoyar la escritura académica. Extraen las frases más importantes de la literatura en trabajos de revisión de los estudiantes. Cada frase clave se corresponde con un artículo de Wikipedia y se califica en una de las cinco categorías concepto abstracto: Investigación de Campo, Tecnología, Sistema, plazo y otros. Este enfoque es adecuado para la escritura de la revisión de literatura; sin embargo, no se toma en cuenta las demás partes del artículo.

**AC6 (Pabitha, Mohanna, Suganthi, & Sivanandhini, 2014): Automatic Question Generation System**

En este artículo los autores proponen un sistema que soluciona los problemas de escritura utilizando el Método de Aprendizaje Supervisado, también denominado método Naïve Bayes. Los autores mencionan que el proceso de automatizar la generación de preguntas consiste en muchas tareas, por ejemplo: Seleccionar el contenido (qué preguntar), el tipo de pregunta (quién, por qué, cómo) y la generación de preguntas reales. El enfoque propuesto abarca solo la generación de preguntas con un método supervisado, el cual no

permite que el sistema aprenda de manera continua y genere nuevas preguntas con la información reciente.

**AC7 (Le, Nguyen, Seta, & Pinkwart, 2014): Automatic question generation for supporting argumentation**

Los autores de este estudio, presentan un enfoque técnico para generar preguntas a petición de los estudiantes durante el proceso de argumentación colaborativa. La contribución de este trabajo radica en la combinación de diferentes tecnologías de procesamiento de lenguaje natural y la explotación de la información semántica para apoyar a los usuarios a desarrollar sus argumentos en una sesión de discusión a través de preguntas a la medida y de diferente tipo. En este estudio se evidencia que en el enfoque planteado no se crea una planificación de las preguntas para obtener posteriormente respuestas que puedan ser utilizadas en párrafos completos.

**AC8 (Le & Huse, 2016): Intelligent Tutoring with Natural Language Support in the Beetle II**

Los autores de este artículo presentan un Sistema de tutoría de diálogo para aceptar entradas de lenguaje sin restricciones y soportar la experimentación con diferentes estrategias de planificación y diálogo. Se demostró que el sistema Beetle II puede utilizarse con éxito como una plataforma para estudiar el impacto que tiene el utilizar diferentes enfoques de tutoría, sin embargo, el sistema no admite una variedad de parámetros que puedan afectar el aprendizaje en sistemas de tutoría inteligente.

**AC9 (Hendel, 2014): Enhancing Writing through Strengthened Executive Function**

En el trabajo de Hendel se mencionan dos estrategias para mejorar la escritura: un método de diagrama de árbol y un método de referencia. Ambos métodos tratan con el flujo de información y son independientes del contenido. Ya existe alguna literatura que indica que estos métodos reducen tres tipos de

ansiedad al escribir y mejoran la organización de la escritura. También hay literatura que propone cuatro medidas específicas de evaluación numérica. Se necesita una evaluación adicional usando estudios de doble enlace que muestren una mejora estadísticamente significativa antes y después. Esta evaluación se debe hacer preferiblemente en el nivel universitario.

Dado un tema de discusión, los estudiantes pueden a veces no proceder a su argumentación. ¿Pueden las preguntas que están semánticamente relacionadas con un tema de debate de ayuda a los estudiantes desarrollar más argumentos?

**AC10 (Kirschner, Eckle-Kohler, & Gurevych, 2015): Linking the Thoughts: Analysis of Argumentation Structures in Scientific Publications**

En este artículo los autores presentan los resultados de un estudio de anotación centrado en el análisis de estructuras de argumentación en publicaciones científicas. Se propone un nuevo esquema de anotación específica cuatro tipos de relaciones argumentativas binarias entre oraciones, resultando en la representación de argumentos como estructuras de grafos pequeños. Proponen una herramienta de anotación que soporta la anotación de dichos gráficos. Esta investigación propone una conexión visual entre argumentos, pero estos podrían verse guiados por expertos en el área.

**AC11 (Gorrostieta & López-López, 2016): Argumentation Identification for Academic Support in Undergraduate Writings**

En este estudio los autores presentan un modelo para guiar la manera en que se aborda la tarea de escribir. También se discute un esquema de anotación para crear un corpus de texto académico con argumentos reconocidos, anotar una muestra y realizar experimentos para identificar automáticamente la argumentación en párrafos, mostrando una eficacia similar a otros enfoques en otros tipos de texto. En este artículo se realiza una investigación inicial; sin embargo, no hay una aplicación de la misma en un sistema o prototipo.

**AC12 (Liu, Lee, & Ding, 2012): Intelligent computer assisted blog writing system**

Los autores de este estudio proponen un sistema inteligente asistido por ordenador para la escritura de blogs. El sistema incluye un módulo de expansión de conceptos, un módulo de generación de texto y un módulo para remplazar el contenido. Uno de los módulos empleados es el módulo de expansión de conceptos que permite ampliar un concepto proporcionado por el usuario y, a su vez el concepto expandido proporciona más información semántica al sistema. En esta propuesta el prototipo del sistema ha demostrado que podría funcionar bien en el dominio de aplicaciones de escritura de blogs, sin embargo, hace falta crear un orden para la información de los conceptos expandidos.

**AC13 (Wilcock & Yamamoto, 2015): Towards Computer-Assisted Language Learning with Robots, Wikipedia and CogInfoCom**

En este estudio, los autores mencionan nuevos enfoques en el que el aprendizaje del lenguaje asistido por robots es reforzado por el uso de canales CogInfoCom para ampliar el compromiso y el interés del alumno. De igual forma, el artículo se centra en el uso de un robot que brinde acceso a la información en Wikipedia para ampliar la gama de temas que el robot puede hablar durante el aprendizaje de idiomas. A pesar de que el robot puede dialogar sobre temas personales aumentando el interés de los estudiantes, este sistema de dialogo no es guiado por un experto, lo que conllevaría a que el sistema no tenga un objetivo comunicacional claro.

**AC14 (Liu, Calvo, & Rus, 2014): Automatic Generation and Ranking of Questions for Critical**

La habilidad de revisión crítica es un aspecto importante de la escritura académica. Las preguntas genéricas del disparador se han utilizado extensamente para apoyar esta actividad. Cuando los estudiantes tienen un tema concreto en mente, las preguntas de activación son menos efectivas si son demasiado generales. En este artículo los autores presentan un sistema basado

en el aprendizaje-a-rango que automáticamente genera preguntas específicas de activación de citas para el apoyo de revisión crítica. Se evaluó el desempeño de los modelos de clasificación de preguntas propuestas y se reportó la calidad de las preguntas generadas. El análisis reveló que las preguntas de explicación y asociación son los tipos de preguntas más frecuentes y que las preguntas de explicación son consideradas como las más valiosas por los estudiantes escritores.

#### **AC15 (Thirugnanasambantham, 2011): An Intelligent System Framework for an Automated Language Tutoring Tool**

En este trabajo, el autor examina el caso de un sistema inteligente de tutorías en idiomas que ha automatizado el proceso de enseñanza, ayudando a aumentar la productividad y a reducir el costo que implica este proceso. Este artículo propone un marco en el que incorporan un sistema de tutoría de lenguaje inteligente basado en un algoritmo de búsqueda de cadenas que extrae patrones de palabras lo que facilita que el sistema proporcione resultados precisos.

El sistema ha demostrado proporcionar una alternativa favorable a la enseñanza de escritura de inglés básico, sin embargo, aprender a escribir es muy importante, pero en muchas ocasiones conlleva demasiado tiempo, algunas investigaciones necesitan ser publicadas rápidamente y no se dispone de este recurso indispensable. Por lo que en estas situaciones sería conveniente una herramienta que facilite y guíe el proceso de escritura en lugar de enseñarlo paso a paso.

#### **AC16 (Hastings, Hughes, Britt, Blaum, & Wallace, 2014): Toward automatic inference of causal structure in student essays**

En este trabajo los autores presentan un método para inferir la estructura causal de ensayos de estudiantes. Utilizan un analizador estándar para derivar dependencias gramaticales del ensayo y los convierten en estados lógicos. A continuación, utilizan un mecanismo de inferencia para identificar conceptos vinculados a los conectores sintácticos por estas dependencias. El estudio indica

que más adelante el método podría proporcionar una retroalimentación explícita que permita a los profesores y estudiantes una mejor comprensión. Este enfoque podría ser utilizado con mayor potencialidad si se lo aplica en un software asistido para la escritura.

**AC17 (Mu, van Aalst, Chan, & Fu, 2014): Automatic Coding of Questioning Patterns in Knowledge-building Discourse**

Los autores describen en este estudio un método propuesto para extraer información sintáctica y semántica antes de la segmentación de los datos en bruto y anotación de acuerdo con un marco de múltiples capas, con lo que se obtienen resultados prometedores que no los utilizan en una aplicación que maneje estos factores para la creación de lineamientos que pueden convertirse en guías para un escritor.

**AC18 (Lochbaum, 1998): A Collaborative Planning Model of Intentional Structure**

En este artículo los autores, proporcionan un modelo computacional para reconocer la estructura intencional y utilizarla en el procesamiento del discurso. En el estudio se menciona que la habilidad de un agente para entender un enunciado depende de su capacidad para relacionar una expresión con el discurso que le precede. El modelo de procesamiento del discurso desarrollado en este artículo proporciona un medio para procesar la comunicación de los agentes. Este enfoque sería mejorado utilizando la información obtenida en una aplicación que guíe al usuario a estructurar correctamente sus ideas.

**AC19 (Trausan-Matu, Dascalu, & Dessus, 2012): Textual Complexity and Discourse Structure in Computer-Supported Collaborative Learning**

En este artículo los autores, muestran un modelo donde se visualiza una estrecha correlación entre la colaboración entre estudiantes en un dialogo y la complejidad textual. Se propone un modelo basado en polifonía, que considera el contenido semántico el cual analiza relaciones entre textos, por ejemplo textos

colaborativos entre estudiantes. Este enfoque puede ser llevado a la práctica en la implementación de sistemas que guíen el proceso de escritura utilizando las tecnologías de Aprendizaje Colaborativo Asistido por Computadora.

**AC20 (Dodigovic, 2013): Identifying off-topic student essays without topic-specific training data**

En este estudio los autores describen un algoritmo que detecta cuándo el ensayo de un alumno está fuera del tema sin requerir un conjunto de ensayos específicos para el tema. Este nuevo sistema es comparable en rendimiento a modelos anteriores que requieren ensayos temáticos específicos para la capacitación, y proporciona información más detallada sobre la forma en que un ensayo se desvía del tema de ensayo solicitado, sin embargo, el sistema no guía al estudiante en el proceso de escritura.

**AC21 (Geithner & Pollastro, 2015): A case based reasoning model for multilingual language generation in dialogues**

En este artículo los autores, proponen mostrar una técnica de Razonamiento Basado en Casos que es fácilmente extensible y adaptable a múltiples dominios y lenguajes, que genera frases coherentes y produce un resultado natural en el contexto de un Agente Conversacional que mantiene un diálogo con el usuario, sin embargo, el proceso de Generación de Lenguaje Natural para un Agente Conversacional traduce algún lenguaje semántico a su forma superficial expresada en lenguaje natural.

**AC22 (Jayalakshmi, Suvarna, Vidya, & Sujatha, 2015): Computer assisted writing system**

En este trabajo los autores, han diseñado e implementado un sistema de escritura asistida por ordenador y el dominio de aplicación es una carta de amor. El sistema incluye módulo de generación de texto, módulo de sustitución de sinónimo y módulo de expresión de símil. Se propone un modelo de generación de texto basado en el modelo de generación de palabras clave y el modelo de

generación de oraciones. El prototipo del sistema ha demostrado que podría funcionar bien en el dominio de la aplicación de una carta de amor y el concepto de esta investigación podría extenderse a otro dominio con una modificación menor. Con la modificación de este sistema se podría conseguir una alternativa que brinde una solución a los problemas de escritura científica.

**AC23 (Roscoe, Allen, Weston, Crossley, & McNamara, 2014):  
Argumentation mining**

Los autores presentan diferentes métodos para ayudar a la minería de argumentación, comenzando con la simple detección de la argumentación y avanzando hacia un análisis más estructural de la argumentación detectada. Diferentes técnicas de última generación en el aprendizaje de máquinas y gramáticas libres de contexto se aplican para resolver los desafíos de la minería de argumentación. También destacan las preguntas fundamentales que se encuentran durante su investigación y analizan diferentes temas sobre la minería de la argumentación. Esta investigación podría ser una de las bases para construir un sistema que permita minimizar los problemas de la escritura científica.

**Conclusiones del estudio del estado del arte**

En esta sección se han propuesto diversas tecnologías para mejorar la escritura. Por ejemplo, existen los denominados Sistemas de Aprendizaje de Lenguaje Asistido por Computadora y los Sistemas de Tutoría Inteligente para Lenguaje; el uso de estas tecnologías mejora la comprensión del lenguaje y la escritura de oraciones, párrafos y documentos. Muchas de estas herramientas están basadas en la Generación del Lenguaje Natural. A esto se le añade la estrategia de Generación Automática de Preguntas que es utilizada por los Tutores Inteligentes y que ayuda a los estudiantes a mejorar su comprensión del texto. A pesar de las diversas propuestas contenidas en cada uno de los artículos científicos expuestos en esta sección, no se han encontrado alternativas de solución concretas al problema de la escritura científica, pues tratan temáticas

sobre la enseñanza de la escritura por medio de tutores inteligentes pero no ofrecen una orientación o una guía en la escritura de artículos a nivel universitario, por lo tanto, en esta investigación se propone el uso de los procesos de la Generación del Lenguaje Natural, para guiar la escritura a través de preguntas. Estas preguntas se desarrollarán con ayuda de la estrategia de Generación Automática de Preguntas y de este modo se creará un modelo que guíe la escritura de un artículo científico. El modelo contemplará un proceso similar al diálogo propuesto por los Sistemas de Tutoría Inteligente, no obstante, no se considera realizar una retroalimentación durante el proceso.

## CAPÍTULO III

### METODOLOGÍA Y MARCO TEÓRICO

En este proyecto de investigación se aborda la escritura de artículos científicos, haciendo referencia a la estructuración de ideas en la formulación del resumen; para lo cual se ha creado un prototipo que facilite a los investigadores el desarrollo de la escritura científica.

La elección apropiada de una metodología orienta el trabajo a través de las diferentes etapas del ciclo de vida del proyecto. En este capítulo se especifica la metodología a ser utilizada, considerando la aplicabilidad de los métodos y técnicas al caso de estudio; además, se expone el marco teórico necesario para sustentar el desarrollo del presente proyecto.

#### **3.1. Metodología**

##### **3.1.1. Design science**

El enfoque metodológico que se seguirá en el presente proyecto es Design Science Methodology for Information Systems and Software Engineering propuesta por Roel J. Wieringa (2014).

La metodología Design Science está orientada al diseño e investigación de artefactos en su contexto, por ejemplo, el diseño e investigación de algoritmos de planificación de rutas basadas en agentes.

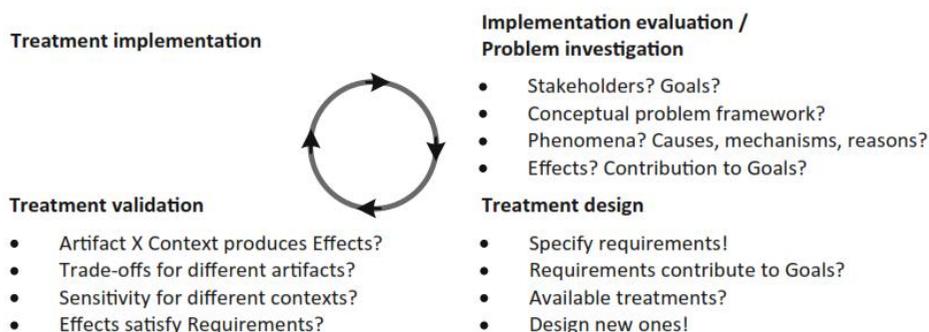
Ésta metodología busca extender los límites de las capacidades humanas y organizacionales creando artefactos nuevos e innovadores. En la metodología, el conocimiento y la comprensión de un dominio del problema y su solución se consiguen en la construcción y aplicación del arte hecho (Hevner, March, Park, & Ram, 2004).

Es fundamentalmente una metodología de resolución de problemas que busca crear innovaciones que definan las ideas, las prácticas, las capacidades técnicas y los productos a través de los cuales el análisis, diseño, implementación, manejo y uso de los sistemas de información puedan ser

efectiva y eficientemente logrados (Denning, 1997), (Tsichritzis, 1998). Tales artefactos no están exentos de leyes naturales o teorías conductuales. Por el contrario, su creación se basa en las teorías existentes que son aplicadas, probadas, modificadas y extendidas a través de la experiencia, la creatividad, la intuición y las capacidades de resolución de problemas del investigador (Markus et al. 2002), (Walls et al., 1992).

Aplicar ésta metodología, nos permite obtener una herramienta que facilite la redacción de artículos científicos por medio del análisis de la información que lo compone; es necesario tomar en cuenta las diferentes guías ya existentes para la redacción de artículos científicos, ya que servirán para el establecimiento de los parámetros adecuados que brinden al redactor, un esquema general de los componentes necesarios para su artículo técnico, siendo importante comprender qué es un diseño de proyecto científico, el cual según indica posee dos componentes esenciales: el objeto de estudio, que permite tener el conocimiento adecuado de las metas del proyecto y el contexto del conocimiento para tener un dominio sobre el proyecto, logrando de esta forma una contribución coherente por medio de la investigación científica (Wieringa, 2014).

La metodología Design Science se estructura de acuerdo con dos grandes ciclos de resolución de problemas: el ciclo de diseño y el ciclo empírico. El primero contiene el ciclo de ingeniería que es un proceso racional de resolución de problemas. El segundo, es una forma racional de contestar a las preguntas del conocimiento científico y está estructurado como una lista de verificación.



**Figura 5 Ciclo de ingeniería**

Fuente: (Wieringa, 2014).

En la figura 5 los signos de interrogación indican preguntas de conocimiento y los signos de exclamación indican problemas de diseño.

El ciclo de Ingeniería define varios términos a ser utilizados en el desarrollo de un proyecto:

- **Tratamiento:** Es la interacción entre el artefacto y el contexto del problema. El investigador diseña no sólo un artefacto, sino que diseña una interacción deseada entre el artefacto y el contexto del problema, destinado a tratar el problema; en este caso el problema es la dificultad que tienen los investigadores al escribir artículos científicos.
- **Artefactos:** Un artefacto es algo creado por la gente para algún propósito práctico. En esta investigación, se construirá un prototipo que brinde una solución al problema antes mencionado.
- **Implementación:** Es la aplicación del tratamiento al contexto original del problema. En este proyecto será la aplicación del prototipo para guiar la redacción de artículos científicos.
- **Validación:** En el ciclo de ingeniería, la validación se realiza antes de la implementación. Consiste en investigar los efectos de la interacción entre un prototipo de un artefacto y un modelo del contexto del problema.

Se seleccionó la metodología Design Science, ya que al enfocarse en Sistemas de Información se adapta en gran medida al proyecto de investigación planteado puesto que el principal objetivo de este estudio es la inserción de un modelo con el fin de mejorar el proceso de escritura en los investigadores universitarios, además ésta metodología permite cubrir cada una de las brechas de investigación que se manifiestan en el desarrollo del presente proyecto desde el planteamiento del problema hasta su validación. La metodología propone 3 tareas, que se describen a continuación:

- Investigación del problema: ¿Qué fenómenos deben mejorarse? ¿Por qué?

En la presente tesis, se ha planteado la problemática sobre la dificultad que tienen los investigadores en la escritura de artículos científicos para lo cual se requiere una entrevista preliminar con investigadores que se encuentren realizando artículos científicos para conocer con mayor profundidad cuales son las dificultades que se presentan durante la escritura y porqué un gran porcentaje de artículos no son publicados en revistas indexadas, para esto la metodología brinda varias guías para la recolección de la información y posteriormente el tratamiento que debe darse a la misma para obtener un modelo que sirva como alternativa de solución al problema.

- Diseño del tratamiento: Diseñe uno o más artefactos que podrían tratar el problema.

Luego de definir e investigar acerca del problema propuesto en la presente tesis, debe darse un tratamiento a la información recolectada para posteriormente crear un modelo que interactúe con los investigadores que en su vida diaria deben enfrentarse a un proceso de escritura con varias limitantes.

- Validación del tratamiento: ¿Estos diseños tratarían el problema?

Validar un artefacto, en este caso un modelo consiste en investigar los efectos de la interacción entre un prototipo y un modelo. En el ciclo de ingeniería, la validación se realiza antes de la implementación.

El objetivo de la validación es predecir cómo un artefacto interactuará con su contexto, sin realmente observar un artefacto implementado en un contexto del mundo real. La investigación de validación es experimental y se realiza generalmente en el laboratorio. En la investigación de validación, exponemos un prototipo de artefacto a varios escenarios presentados por un modelo del contexto, para ver cómo responde.

Por otra parte, este proyecto de investigación tendrá un nivel analítico, el cual consiste en el análisis final de los datos, la forma en que serán procesados,

las pruebas a los que serán sometidos y los criterios sobre los que serán interpretados (Universidad del Rosario, 2003).

Según el nivel de conocimiento que se adquiere en la investigación, el proyecto es de tipo investigación descriptiva ya que busca relatar el conocimiento adquirido y socializarlo a la comunidad científica (Kirsch & Sullivan, 1992).

### **3.1.2. Amenazas a la validez**

- Un diálogo puede no ser adaptado por completo a todas las áreas de investigación.
- Es posible que el modelo no cumpla plenamente con los objetivos de las partes interesadas.
- El número de estudios seleccionados como base para el presente proyecto, pueden ser insuficientes para realizar una investigación más amplia.

## **3.2. Marco teórico**

### **3.2.1. Fundamentación Filosófica**

Cuando aparecieron las calculadoras automáticas y las posteriores teorías computacionales, de la mano de la creación de la máquina de Turing, se creyó que estas máquinas serían capaces de superar fácilmente la inteligencia humana. Ahora, sin embargo, se las considera máquinas tontas por la dificultad que tienen al realizar tareas que son normales para las personas, como generar ideas, comunicarse con una persona o trabajar en grupo. Una computadora no puede hacer nada; a menos que sea programada con algoritmos determinísticos; o diseñada para que aprenda, con las redes neuronales.

Einstein decía “Si no se lo puedes explicar a un niño de 6 años, entonces tú tampoco lo entiendes”. Esta lección del conocido científico puede ser estudiada, y se puede decir que la comunicación tiene un complejo y amplio significado. La lección de Einstein puede deberse a que la situación hipotética propuesta por él, o el experimento mental como él solía llamar a las situaciones imaginarias, cuentan con un factor crucial, la ignorancia de un niño de 6 años. La ignorancia no es más que el desconocimiento del contexto de un tema dado.

Una computadora es un mejor ejemplo de un individuo, ente, agente o como se quiera llamarlo; que cuenta con mucha ignorancia o desconocimiento de un contexto determinado, es por esta razón que para instruir a un computador es necesario explicar todos y cada uno de los detalles de un proceso, con los conocidos algoritmos, que no son más que flujos o formas de realizar una tarea paso a paso, en las cuales existen bifurcaciones o situaciones en donde es necesario tomar una decisión.

Toda tarea que puede realizar una computadora, es posible que la realice un humano y no es así de forma contraria. Las consideraciones a tomar en cuenta al momento de que un humano realice una tarea que normalmente se la delegaría a una computadora es que se la delega a una computadora porque el proceso es demasiado extenso o necesita de mucha precisión. La ventaja de que un proceso haya sido automatizado es que ha sido totalmente entendido, como es necesario, para que un niño de 6 años lo entienda, y más aún, para que un computador este instruido para ejecutarlo. Si no se ha entendido totalmente se ha encontrado un patrón con resultados positivos a la hora de realizar el proceso.

Dentro de los procesos complicados de automatizar está la escritura, aunque se ha reportado una gran cantidad de investigaciones a lo largo de los últimos años. La Generación de Lenguaje Natural es una rama del Procesamiento del Lenguaje Natural que trata acerca de los mecanismos que se han utilizado a la hora de que las computadoras generen lenguaje entendible para el humano.

La escritura es un proceso que necesita un fuerte trabajo cognitivo, son muchas las tareas necesarias para realizar un escrito. El proceso se complica aún más cuando un escrito debe ser revisado y aprobado por algún motivo en particular. En este estudio se propone utilizar los mecanismos y soluciones encontradas en la informática para tratar el proceso de la escritura.

### 3.2.2. Hipótesis

Los procesos de la generación del lenguaje natural y los sistemas de tutoría inteligente pueden ser utilizados para analizar y plantear soluciones a la dificultad en la escritura científica y la diferenciación de sus sub-tareas.

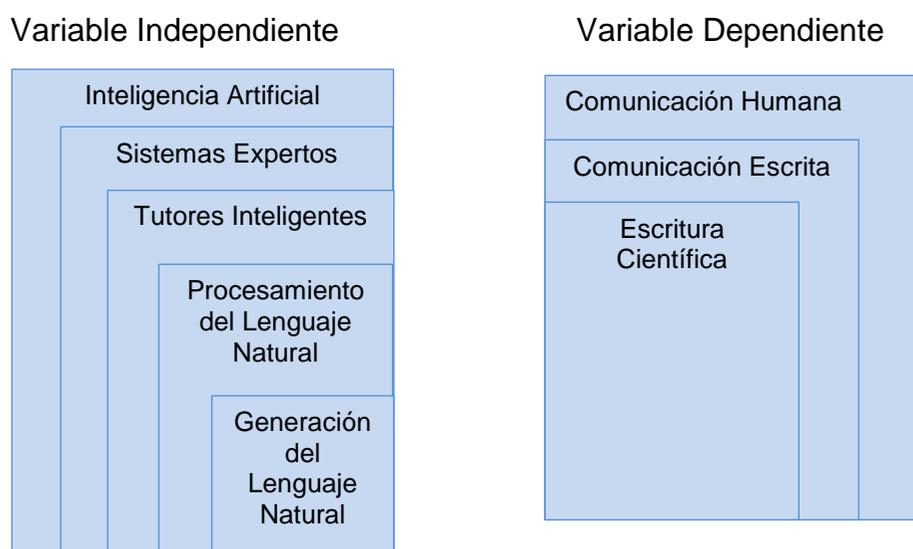
#### Señalamiento de Variables:

- **Variable Independiente:** Los procesos de la generación del lenguaje natural y los sistemas de tutoría inteligente.
- **Variable Dependiente:** La dificultad en la escritura científica y la diferenciación de sus sub-tareas.

Para el cumplimiento de la hipótesis planteada para este trabajo de investigación, se desarrollará un modelo que es detallado con cada una de sus fases en los siguientes capítulos de la presente tesis.

### 3.2.3. Red de Categorías

Con la finalidad de buscar la congruencia en la fundamentación teórica de la presente tesis, es fundamental establecer una red de categorías que contribuyan en la explicación y entendimiento científico del tema en estudio; la red de categorías se muestra a continuación en la Figura 6.



**Figura 6 Red de Categorías de las Variables de Investigación**

### **3.2.4. Fundamentación Científica De La Variable Independiente**

#### **Inteligencia Artificial**

Se puede definir a la Inteligencia Artificial como una ciencia que tiene como objetivo el diseño y construcción de máquinas capaces de imitar el comportamiento inteligente de las personas. Una rama especializada de la Informática que investiga y produce razonamiento por medio de máquinas automáticas y que pretende fabricar artefactos dotados de la capacidad de pensar (Álvarez Munárriz, 1994).

Para que la inteligencia artificial tenga éxito, necesitamos dos cosas: la inteligencia y un artefacto. El ordenador ha sido el artefacto de elección. La moderna computadora electrónica digital fue inventada de forma independiente y casi simultáneamente por científicos de tres países asediados en la Segunda Guerra Mundial. El primer ordenador operativo fue el electromecánico Heath Robinson, construido en 1940 por el equipo de Alan Turing con un único propósito: descifrar mensajes alemanes (Norvig & Russell, 2010).

El trabajo en Inteligencia Artificial ha sido pionero en muchas ideas que han hecho su camino de regreso a la informática convencional, incluyendo el tiempo compartido, intérpretes interactivos, computadoras personales con ventanas y ratones, tipo de datos de lista, gestión de almacenamiento automático y conceptos clave de programación simbólica, funcional, declarativa y orientada a objetos (Norvig & Russell, 2010).

La lingüística moderna y la Inteligencia Artificial "nacieron" casi al mismo tiempo, y crecieron juntas, intersectándose en un campo híbrido llamado lingüística computacional o procesamiento del lenguaje natural. Comprender el lenguaje requiere una comprensión del tema y el contexto, no sólo una comprensión de la estructura de las oraciones. Gran parte del trabajo inicial en la representación del conocimiento (el estudio de cómo poner el conocimiento en una forma con la que un ordenador puede razonar) estaba ligado al lenguaje, que a su vez estaba conectada a décadas de trabajo sobre el análisis filosófico Del lenguaje (Norvig & Russell, 2010).

## **Sistemas Expertos**

Los sistemas expertos se orientan a la resolución de problemas complejos que de por sí necesitan la pericia de un experto para poder superarlos (Amador Hidalgo, 1996).

Por lo general los sistemas expertos suelen incorporar un cuerpo de conocimiento (capacidad de inferir, experiencia, memoria, certidumbre, etc.) considerable sobre una determinada área del saber que permite, conjugado con otros elementos, reproducir la conducta seguida por cualquier experto en dicha área y, más concretamente, en su actuación corriente a la hora de solventar los problemas que son propios de su actividad (Amador Hidalgo, 1996).

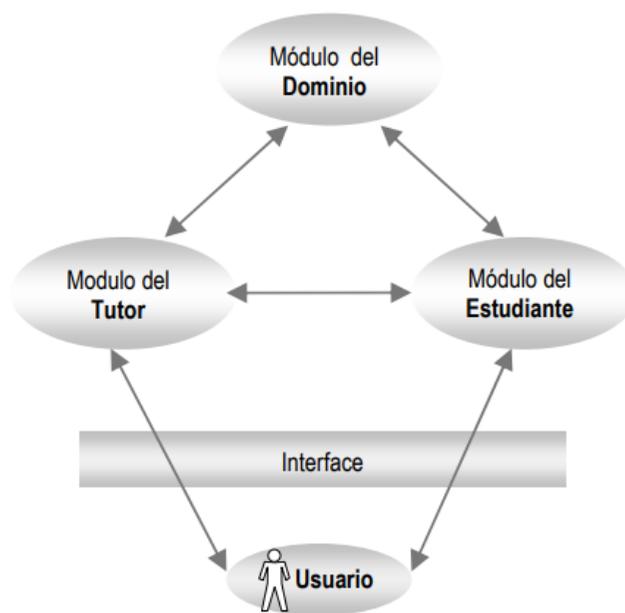
## **Tutores Inteligentes**

Un tutor inteligente es un sistema que usa métodos de inteligencia artificial para plasmar el conocimiento e interactuar con los estudiantes (VanLehn, 1988 en Cataldi & Lage, 2010).

Wolf (citado en Cataldi & Lage, 2010) expone a los Sistemas de Tutoría Inteligente como: “sistemas que modelan la enseñanza, el aprendizaje, la comunicación y el dominio del conocimiento del especialista y el entendimiento del estudiante sobre ese dominio”.

Un sistema que incorpora técnicas de Inteligencia Artificial a fin de crear un ambiente que considere los diversos estilos cognitivos de los alumnos que utilizan el programa (Giraffa, 1997 en Cataldi & Lage, 2010).

Los sistemas de tutoría inteligente son capaces de proporcionar un enfoque personalizado al aprendizaje asumiendo el rol de un verdadero experto que orienta en el proceso de aprendizaje de acuerdo a las necesidades específicas de cada alumno.



**Figura 7 Interacción de los Módulos de un STI**

Fuente: (Cataldi & Lage, 2010).

La arquitectura de un Sistema de Tutoría Inteligente contiene varios modelos funcionales: modelo de dominio o experto, modelo de conocimiento, modelo de aprendizaje y modelo de tutoría o interfaz de usuario (Cataldi & Lage, 2010).

- Modelo de dominio o experto: Contiene todo el conocimiento experto que debe ser enseñado por el sistema.
- Modelo de aprendizaje: Contienen información relevante del alumno, como su nivel previo de conocimiento lingüístico, preferencia de estilo de aprendizaje, metas o interés.
- Modelo de tutoría o interfaz de usuario: Se compone de reglas que definen cómo un sistema debe actuar durante el tiempo de ejecución, teniendo en cuenta el modelo de aprendizaje.
- Modelo de interfaz de usuario: Se refiere al entorno gráfico a través del cual el contenido educativo se entrega al alumno.

## **Procesamiento Del Lenguaje Natural**

Es un área de investigación y aplicación que explora cómo las computadoras se pueden utilizar para entender y manipular el lenguaje natural del lenguaje o el habla común (Chowdhury, 2003).

Los fundamentos del Procesamiento del Lenguaje Natural se encuentran en una serie de disciplinas, como la informática y las ciencias de la información, lingüística, matemáticas, ingeniería eléctrica y electrónica, inteligencia artificial y robótica, y la psicología (Chowdhury, 2003).

Las aplicaciones del PNL incluyen una serie de campos de estudio, tales como generación del lenguaje natural, respuestas a preguntas, traducción automática, síntesis de discurso, procesamiento y resumen de texto en lenguaje natural, interfaces de usuario, reconocimiento de voz, inteligencia artificial y sistemas expertos (Chowdhury, 2003). El estudio del lenguaje natural está estructurado por 3 niveles de análisis:

### **Análisis sintáctico**

La Sintáctica corresponde al análisis de la relación existente entre los distintos símbolos o signos del lenguaje (Vega Mora, 2008).

La Sintaxis especifica las reglas según las cuales una expresión, por ejemplo, una oración, está bien formada (Van Dijk & Mayoral, 1987).

Se denomina así al servicio que una palabra desempeña en otra o, de forma más precisa, las relaciones de combinación (Dik, Valverde, & Mingorance, 1981).

Se refiere al lugar que ocupa dentro de la oración, estudia la estructura de la oración y el modo en que se enlazan las palabras para formar oraciones con sentido (Dik, Valverde, & Mingorance, 1981).

La importancia de la sintaxis en la comunicación se basa en que orienta la construcción adecuada de oraciones, dando como resultado una expresión oral coherente (Dik, Valverde, & Mingorance, 1981).

### **Análisis semántico**

La semántica especifica las reglas de acuerdo con las cuales dicha expresión es portadora de significado (Van Dijk & Mayoral, 1987).

La Semántica es el proceso de significación mediante el lenguaje (Dik, Valverde, & Mingorance, 1981).

Es el estudio de la relación entre los signos y su significado. Se conoce como semántica al estudio del significado de los signos lingüísticos y de sus combinaciones. Está vinculada al significado, sentido e interpretación de palabras, expresiones o símbolos (Dik, Valverde, & Mingorance, 1981).

La semántica es la ciencia de los significados, donde el producto es un sistema en el que se pueden originar y expresar diferentes lenguajes, dando sentido y significado al objeto (Dik, Valverde, & Mingorance, 1981).

### **Análisis pragmático**

El análisis pragmático se ocupa de la formulación de las reglas según las cuales un acto verbal es apropiado en relación a un contexto (Van Dijk & Mayoral, 1987).

Pragmático es relativo a la práctica o la realización de las acciones. El análisis pragmático es una disciplina que estudia el lenguaje en relación al contexto donde se desarrolla la idea, es decir, las oraciones producen una aceptación semántica pero su significado e interpretación depende del contenido y el contexto lingüístico, ya que una misma oración puede tener varios sentidos en diferentes contextos (Dik, Valverde, & Mingorance, 1981).

El análisis pragmático añade información adicional al análisis del significado de la frase en función del contexto donde parece. Se trata de uno de los niveles de análisis más complejos, la finalidad de la cual es incorporar al análisis semántico la aportación significativa que pueden hacer los participantes, la evolución del discurso o información presupuesta (Dik, Valverde, & Mingorance, 1981).

## **Generación del Lenguaje Natural**

La Generación del Lenguaje Natural (GLN) es el subcampo de la inteligencia artificial y la lingüística computacional que se centra en los sistemas informáticos que pueden producir textos comprensibles en inglés u otras lenguas humanas. Normalmente, partiendo de alguna representación no lingüística de la información como entrada, los sistemas GLN utilizan el conocimiento sobre el lenguaje y el dominio de la aplicación para producir automáticamente documentos, informes, explicaciones, mensajes de ayuda y otros tipos de textos. GLN es a la vez una fascinante área de investigación y una tecnología emergente con muchas aplicaciones del mundo real (Reiter & Dale, 2000).

Como un área de investigación, GLN trae una perspectiva única sobre cuestiones fundamentales en la inteligencia artificial, la ciencia cognitiva y la interacción hombre-hombre. Éstas incluyen preguntas tales como ¿Cómo deben ser representados y razonados los conocimientos lingüísticos y de dominio?, ¿Qué se debe hacer para que un texto esté bien escrito y cómo se comunica mejor la información entre la máquina y el ser humano?

Desde una perspectiva práctica, la tecnología GLN es capaz de automatizar parcialmente la creación rutinaria de documentos, eliminando gran parte del trabajo penoso asociado con tales tareas. A más largo plazo, GLN también es probable que desempeñe un papel importante en las interfaces hombre-computadora y permitirá una interacción mucho más rica con las máquinas que es posible hoy en día (Reiter & Dale, 2000).

## **Fases de la Generación del Lenguaje Natural**

### **Macroplanificación**

La primera fase es la macroplanificación o planificación del documento. El planificador de documentos es responsable de decidir qué información debe comunicarse y determinar cómo debe estructurarse esta información para su presentación (tarea de estructuración de documentos). En muchas aplicaciones, el planificador de documentos es el componente más importante del sistema de Generación de Lenguaje Natural (Reiter & Dale, 2000).

- **Determinación del contenido:** es la tarea de decidir qué trozos de información, o, mensajes, deben ser incluidos en el texto generado (Reiter & Dale, 2000).
- **Estructura de Documentos:** es la tarea de construir una estructura que contiene los mensajes seleccionados en la determinación de contenido. Esto requiere agrupar los mensajes que se van a transmitir. La estructuración de documentos también puede especificar el orden en que aparecen los mensajes y las relaciones de discurso que se mantienen entre varios mensajes o grupos de mensajes (Reiter & Dale, 2000).

### **Microplanificación**

La tarea del componente de microplanificación es tomar tal plan de documento y refinarlo para producir una especificación de texto más detallada que se puede pasar al componente de realización de superficie, que producirá un texto de superficie correspondiente. El plan de documentos deja abierta una serie de decisiones sobre la eventual forma del texto en el documento que se va a generar (Reiter & Dale, 2000). Estas decisiones más finas se toman por el componente de microplanificación. En la arquitectura del sistema de GLN, el microplanificador se ocupa de:

- **Lexicalización:** Elección de las palabras particulares, construcciones sintácticas y anotaciones de marcado utilizadas para comunicar la información codificada en el plan de documento.
- **Agregación:** Decidir cuánta información debe ser comunicada en cada una de las oraciones del documento.
- **Generación de expresiones referente.** Determinar qué frases se deben utilizar para identificar entidades de dominio particulares para el usuario.

## Realización

La función del componente realización consiste en tomar la especificación de texto producida por el microplanificador y convertirla en texto. Esto implica recorrer la especificación de texto, típicamente de arriba a abajo, de izquierda a derecha, y producir una salida apropiada de los nodos encontrados durante este recorrido. La información en los nodos proporciona una especificación de lo que el realizador necesita hacer. En un proceso de realización lingüística, las especificaciones de frase pueden especificar unidades lingüísticas en varios niveles de abstracción. La realización lingüística es el área más bien establecida de GLN. Como resultado de esto, una serie de sistemas existentes para llevar a cabo esta etapa de procesamiento están disponibles en el dominio público (Reiter & Dale, 2000).

## Teoría de la Estructura Retórica

La Teoría de la Estructura Retórica (RST) es una teoría de la estructura del texto que se está extendiendo para servir de base teórica para la planificación del texto computacional (Reiter & Dale, 2000). Las estructuras de texto en RST son jerárquicas, construidas sobre pequeños patrones llamados esquemas. Los esquemas que componen la jerarquía estructural de un texto describen las funciones de las partes más que sus características de forma (Reiter & Dale, 2000). Las relaciones entre partes de texto, comparables a las relaciones conjuntivas, son una parte prominente de la maquinaria de definición de RST. Un trabajo reciente ha puesto a RST en una nueva base de definición (Reiter & Dale, 2000).

Algunas de las relaciones comunes del discurso que se utilizan con frecuencia son las siguientes:

- **Contraste:** Un mensaje o grupo de mensajes proporciona información de contraste a la proporcionada en otro mensaje o grupo de mensajes.
- **Elaboración:** Un mensaje o grupo de mensajes explica con detalle la información de otro mensaje o grupo de mensajes.

### **3.2.5. Fundamentación Científica de la Variable Dependiente**

#### **Comunicación humana**

La comunicación humana no es otra cosa que el intercambio de mensajes entre un emisor y un receptor. En este intercambio de mensajes se presentan tres variantes como son la comunicación intrapersonal, interpersonal y de masas. Si bien es cierto que existen muchos tipos de lenguaje y formas de él, el de naturaleza humana se caracteriza porque su naturaleza es arbitraria, posee estructura de significante, es de carácter lineal y posee la condición de doble articulación (Bolaños, 1996).

La comunicación humana se produce por medio de mensajes, que el ser humano emite en dirección con los demás. Estos mensajes constan de ciertas unidades o elementos, portadores de sentido y significado, lo que le permite la comprensibilidad y comunicabilidad individual y social (Bolaños, 1996).

#### **Comunicación escrita**

La comunicación escrita es un modo de transmisión en el que el emisor elabora su mensaje sin necesidad de coincidir espacial y temporalmente con su receptor al momento de hacerlo (Renkema, 1999).

La comunicación escrita codifica con precisión el lenguaje hablado, en este sentido, hay un lazo entre oralidad y escritura, pero la comunicación escrita posee rasgos particulares con respecto al uso del lenguaje (Renkema, 1999).

#### **Escritura científica**

La propiedad principal de la escritura científica es la claridad. Esencialmente, la claridad debería ser una característica de todo tipo de comunicaciones; sin embargo, cuando se expone algo por primera vez, la claridad es fundamental. La mayoría de los artículos científicos publicados en revistas de investigación primarias se aceptan para su publicación justamente porque aportan auténticos conocimientos científicos nuevos. Por ello, es necesario exigir una claridad absoluta en la escritura científica (Day, 2005).

La escritura científica es la transferencia de una señal clara al receptor. Las palabras de esa señal deben ser tan claras, sencillas y ordenadas como sea posible. La redacción científica no tiene necesidad de adornos ni cabida para ellos. Es muy probable que las metáforas, los símiles y las expresiones idiomáticas inciten a confusión, por lo que pocas veces deben utilizarse al escribir artículos de investigación (Day, 2005).

En la escritura científica se dice: “El mejor lenguaje es el que transmite el sentido con el menor número posible de palabras”. Los juegos literarios y las metáforas hacen que la atención se desvíe del contenido al estilo. Deben usarse rara vez, si acaso se usan, en la redacción científica (Day, 2005).

## CAPÍTULO IV

### DISEÑO Y DESARROLLO

#### 4.1. DISEÑO

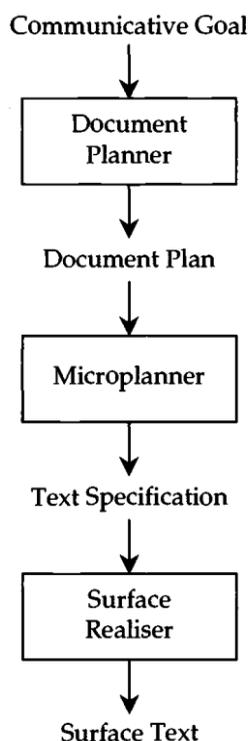
Conforme a los artículos científicos que se obtuvieron para el estudio del estado del arte, seguido de un análisis riguroso se estableció que los mismos servirán como base para la construcción del modelo, ya que todos los artículos manejan una estructura similar y una temática en común respecto a las variables dependientes e independientes determinadas en este estudio. Cabe recalcar que únicamente se analizaron los resúmenes de cada artículo científico, en los cuales se observó su estructura y organización. Para esto, además de los dos investigadores involucrados en este proyecto, se contó con el aporte de un investigador experimentado el cual brindó su conocimiento en el análisis y tratamiento de la información. El modelo es adaptado a las recomendaciones del investigador experimentado partiendo de la necesidad de publicar sus hallazgos científicos.

##### 4.1.1. Diseño del modelo

Dentro del diseño del modelo es necesario conocer la teoría de la Generación del Lenguaje Natural, la cual describe algunas fases principales para el proceso de escritura. Cada una de estas fases será utilizada para la solución del problema de escritura, siendo el caso específico de aplicación la redacción de artículos científicos. De igual forma, se hará uso del análisis de la estructura de los artículos seleccionados (fuente de conocimiento de la que se nutre el prototipo), mismos que permitirán determinar tanto el objetivo comunicativo, como el contexto del discurso.

Asimismo, guiados por el marco conceptual de la Generación del Lenguaje Natural, para cada uno de los artículos se realizaron las fases de estructuración del documento (Macroplanificación) y estructuración de oraciones

(Microplanificación), para lo cual se llevó a cabo un tratamiento exhaustivo de las características de los artículos científicos, tales como estructura, modelo de dominio y su estructuración retórica.

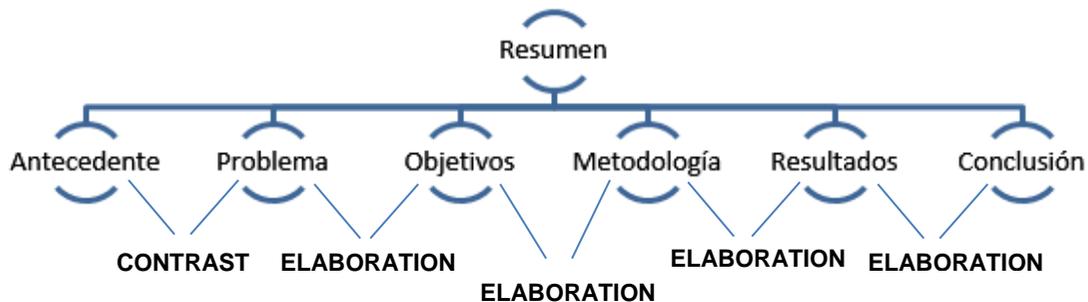


**Figura 8 Arquitectura de un Sistema de GLN**

Fuente: (Reiter & Dale, 2000).

### **Macroplanificación**

En la fase de Macroplanificación, la información es tomada de acuerdo a los lineamientos de la escritura de Artículos Científicos: Resumen, Introducción, Metodología, Resultado y Conclusión, con esta información cumplimos el objetivo de definir la estructura del modelo propuesto obteniendo como resultado la entrada inicial para esta fase y, adicionalmente, como salida de la fase de Macroplanificación se obtuvo un árbol que expone los mensajes a transmitir como nodos terminales y a su vez el orden en el que estos deben ser presentados en el texto final (ver Figura 9).



**Figura 9 Mensajes a transmitir en su respectivo orden**

El orden de los mensajes transmitidos adquiere importancia desde el punto de vista del análisis pragmático. Donde el texto anterior proporciona información al texto sucesor.

### **Microplanificación**

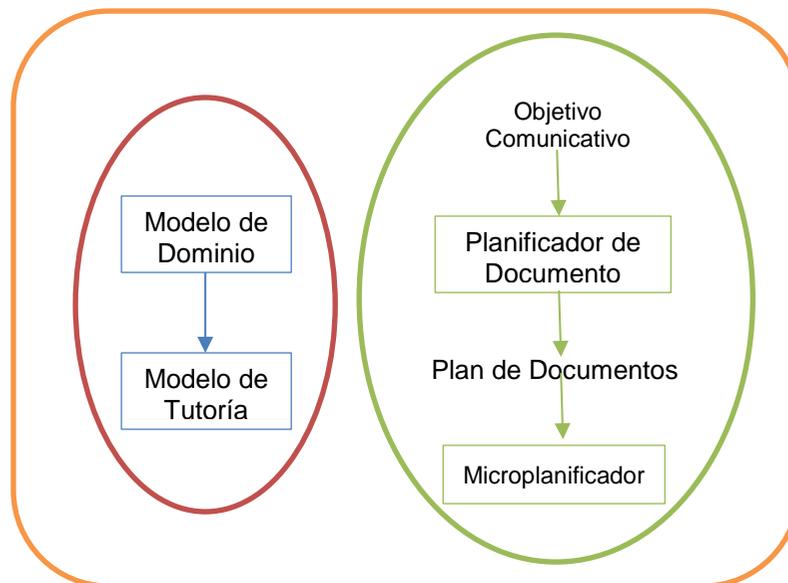
En esta fase se determinaron las palabras acordes a las relaciones entre los mensajes descritos en la Macroplanificación, las cuales fueron extraídas de los artículos estudiados. Por ejemplo: la relación de contraste entre el antecedente y el problema se traduce a la expresión “Sin embargo”.

Además, fue necesario determinar las entidades presentes en el dominio de estudio. Por ejemplo, se encontraron las expresiones “habilidades de escritura”, “sistemas inteligentes”, “minería de argumentación”, “transiciones del discurso” las cuales responden a las variables dependientes e independientes descritas normalmente en las hipótesis de los estudios.

Otras de las entidades encontradas fueron “niños de escuela”, “estudiantes de universidad”, las cuales responden a los involucrados en los estudios y otras entidades como “Inteligencia artificial”, “Tutores inteligentes”, “Generación del Lenguaje Natural” que representan la temática en la que se enmarca el artículo científico.

Con lo descrito en las fases de Macroplanificación y Microplanificación se procede al diseño del siguiente módulo de la propuesta planteada en este

estudio. El cual corresponde al Sistema Inteligente de Tutoría descrito a continuación.



**Figura 10 Combinación de arquitecturas**

Para el modelo del dominio cabe hacer una diferenciación entre el modelo del dominio de la generación del lenguaje natural y el modelo del dominio de los tutores inteligentes. En el primer caso el modelo del dominio tiene como objetivo identificar las características de los artículos científicos para la futura creación de los mismos. Por otro lado, en el segundo caso el modelo del dominio tiene como objetivo identificar los conceptos, reglas y estrategias de resolución de problemas dentro de la creación de artículos científicos.

Sin embargo, el modelo del dominio del sistema de tutoría inteligente o conocimiento experto está alineado al modelo del dominio de la Generación del lenguaje Natural por lo que se tomó sus características para intentar que, al finalizar la tutoría, el usuario aprendiz construya un resumen acotado a lo indicado en el modelo de dominio de la Generación del Lenguaje Natural.

Pese a que es necesario el modelo del estudiante para el modelo del tutor. No fue construido el modelo del estudiante, pero se tomaron ciertas

consideraciones en función del estudiante, como las costumbres que tiene al momento de escribir un artículo. Con esta información se obtuvo el modelo del tutor que se traduce al flujo normal del proceso de la propuesta de este estudio.

Como sustento al diseño del modelo se ha realizado un análisis del contenido de los resúmenes de 18 artículos científicos detallados en la tabla 1.

Para determinar el contenido esencial del resumen se hicieron preguntas a investigadores de la universidad y fue necesaria la fundamentación teórica con libros y artículos relacionados, obteniendo las siguientes secciones: Problema, Antecedente del problema, Objetivo de la investigación, Metodología, Resultados y Conclusiones/Discusión/Implicaciones.

De los artículos analizados se obtuvo la siguiente tabla que describe el orden que se encontró con mayor frecuencia las secciones del resumen.

**Tabla 4**

**Frecuencia de movimientos observados en el resumen de cada artículo**

Antecedente	Problema	Objetivo	Metodología	Resultado	Conclusión
12	0	3	0	2	1
1	7	6	2	0	1
1	2	5	3	3	4
1	0	2	4	5	2
0	1	0	3	1	4
0	0	0	0	3	2
0	0	0	1	1	0
0	0	0	0	1	0

En la tabla 4, se describe el número de artículos científicos que coincidieron con el orden propuesto:

- 12 artículos exponen en primer lugar su Antecedente, 7 artículos tienen su problema en segundo lugar, 5 artículos continúan con su

objetivo y seguido a esto 4 con la metodología, 1 con el resultado y 2 con la conclusión. Cabe mencionar que el orden descrito se toma en referencia al modelo propuesto en este proyecto, es decir, que existen variaciones en el orden respecto a otros artículos en lo que, por ejemplo, se tienen únicamente las secciones de metodología y resultados o, problema, metodología, resultado y conclusión.

Como se puede observar el patrón más encontrado es: antecedente, problema, objetivo, resultado y conclusión, con lo cual se ajusta al modelo encontrado para tener un resumen completo, estableciendo un patrón de similar orden a excepción de que en cuarto lugar se recomienda escribir la metodología como se muestra en el cuadro de color naranja de la tabla anterior. Cabe recalcar que en la mayoría de artículos estudiados no se encontró la sección de metodología en la que se describe de forma breve los pasos que siguieron los investigadores para llegar a la solución del problema planteado.

Una vez definida la estructura y el meta-contenido del resumen se procede a definir las expresiones de referencia. Se define el meta-contenido debido a que, si se define únicamente el contenido, el sistema será capaz tan solo de producir texto de una investigación en particular. Para la definición de las expresiones de referencia se realizará algo similar al meta-contenido.

Para la definición de las expresiones de referencia se realizará un análisis de las propiedades de las secciones del resumen o movimientos retóricos. A continuación, se definirán entidades que analicen los conceptos que se encuentran en las investigaciones. Para lo cual se caracterizarán las secciones retóricas.

Los resultados del modelo sirven de base para la elaboración del diálogo. El orden del modelo se refleja en las preguntas planteadas en el diálogo. Además, se toma en cuenta las meta-entidades encontradas para sugerir al usuario que las tome en cuenta en las respuestas a las preguntas. Quedando el diálogo de la siguiente manera.

1. ¿Indique cuál fue el objeto de estudio sobre el que se desea solucionar o mejorar algo?

Respuesta = variableObjetoEstudio.

2. ¿Seleccione cuál fue la principal herramienta utilizada en el estudio 1, 2,...,7 o ingrese cual es el recurso si no es ninguno?

- a. Una herramienta
- b. Una teoría
- c. Una metodología
- d. Una técnica
- e. Un método
- f. Un modelo
- g. Información

Respuesta = variableRecurso

3. Escriba el nombre o descripción de variableRecurso utilizada.

Respuesta = variableRecursoElegido

4. Escriba una ciencia o campo de estudio que abarque el/la variableRecursoElegido. Esta ciencia o tema, de preferencia, debe ser bien conocido o debe ser de importancia para sus futuros lectores.

Respuesta = variableCienciaOCampo

5. Escriba un hecho de actualidad reconocido por la ciencia que relacione variableCienciaOCampo y variableObjetoEstudio, que muestre la importancia de relacionarlas (antecedente del estudio).

Respuesta = variableAntecedente

6. Describa el hecho más importante reconocido por la ciencia en el que se mencione: el contexto en el que se desenvuelve el fenómeno de estudio, el mal uso de variableRecursoElegido en variableObjetoEstudio o la

inutilización de variableRecursoElegido en variableObjetoEstudio (problema a resolver). Mencione variableCienciaOCampo o variableRecursoElegido y variableObjetoEstudio.

Respuesta = variableProblema

7. Dado que variableProblema. Describa la propuesta de solución (objetivo del estudio), relacione variableRecursoElegido y variableObjetoEstudio.

Respuesta = variableObjetivo

8. Describa técnicamente como utilizó o desarrollo variableRecursoElegido para lograr el cumplimiento del objetivo. Especifique que tipo de efecto realiza la variableRecurso.

Respuesta = variableMetodología

9. Describa como obtuvo los resultados y cuáles son los más relevantes. Realice una afirmación en la que se utilice variableRecursoElegido y variableObjetoEstudio. Y el efecto obtenido.

Respuesta = variableResultados

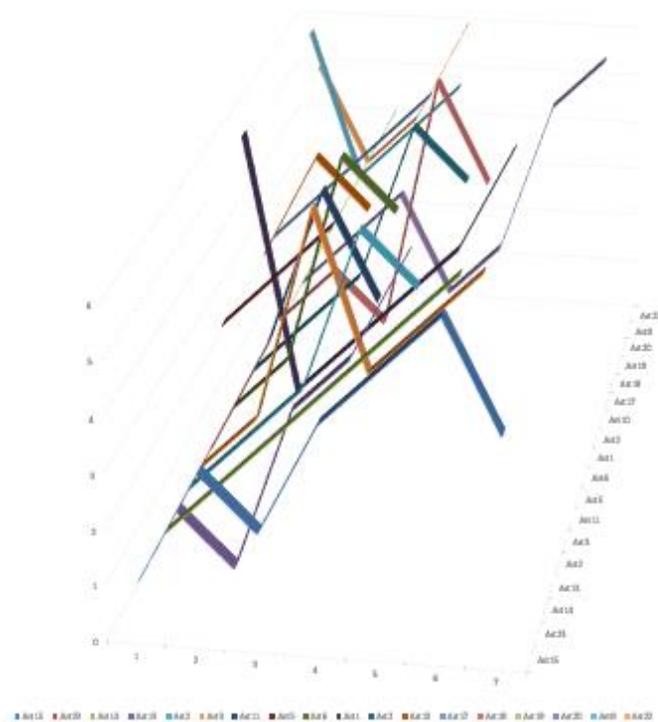
10. Escriba las conclusiones en función del cumplimiento del objetivo basado en los resultados.

Respuesta = variableConclusiones

Como se puede observar el dialogo tiene variables las cuales rellenan las preguntas que les suceden.

Por otro lado, en el siguiente gráfico se muestra el orden en que los artículos tienen estructurados las secciones antes mencionadas. En el eje de las "X" está el orden en que se encuentran las diferentes secciones en el resumen. En el eje de las "Y" están, en el orden que se espera encontrar las diferentes secciones como tal, de esta forma, 1. Corresponde al antecedente, 2. al

problema, 3. objetivo, 4. metodología, 5. resultados, 6 conclusiones/discusión/implicaciones.



**Figura 11 Orden encontrado según la estructura de los artículos**

Adicionalmente, para determinar el número de palabras aproximado que debe tener un resumen, se realizó un conteo del total de palabras contenidas en los resúmenes de los artículos científicos. Obteniendo que el total de caracteres encontrados en promedio fue de 1286, el mínimo de 447 y el máximo de 2691 (ver tabla 5).

**Tabla 5**

**Número de caracteres encontrado**

	Antecedente	Prob.	Obj.	Metodología	Resultado	Conclusión	Suma
<b>Prom.</b>	227	200	161	314	150	234	1286
<b>Mín.</b>	82	73	85	58	44	135	477
<b>Máx.</b>	455	453	259	748	285	491	2691

## 4.2. DESARROLLO

### 4.2.1. Descripción de la técnica

Para plasmar los hallazgos encontrados, conforme a los patrones lingüísticos que se observaron y que sirvieron para plantear el modelo del resumen de los artículos científicos, se realizó un prototipo cuyo fin es reproducir el modelo que se ha diseñado y con el cual los investigadores pueden escribir el resumen de un artículo científico siguiendo las pautas que propone el prototipo.

Estos son los pasos para escribir el abstract de un artículo

Objeto de estudio

Recurso utilizado

1. ¿Indique cuál fue el objeto de estudio sobre el que se desea solucionar o mejorar algo?

Antecedente

Problema

Objetivo

Método

Resultados

Conclusiones

Anterior Siguiete

Texto Final

Ejemplos

Texto recomen...

**Figura 12 Interfaz del prototipo**

El prototipo es un híbrido entre un generador de lenguaje natural y un Sistema Inteligente de Tutoría. El prototipo guía al investigador en la escritura, a través de preguntas enmarcadas en el diálogo al estilo de un tutor humano siguiendo las características del Sistema de Tutoría Inteligente y teniendo en

cuenta la generación de preguntas basado en plantillas. Las preguntas fueron analizadas por los expertos obteniendo así un flujo más natural en las preguntas.

Las preguntas generadas se muestran en la siguiente figura:

Estos son los pasos para escribir el abstract de un artículo

1. ¿Indique cuál fue el objeto de estudio sobre el que se desea solucionar o mejorar algo?

Objeto de estudio

Recurso utilizado

Antecedente

Problema

Objetivo

Método

Resultados

Conclusiones

Anterior Siguiete

Texto recomen...

Ejemplos

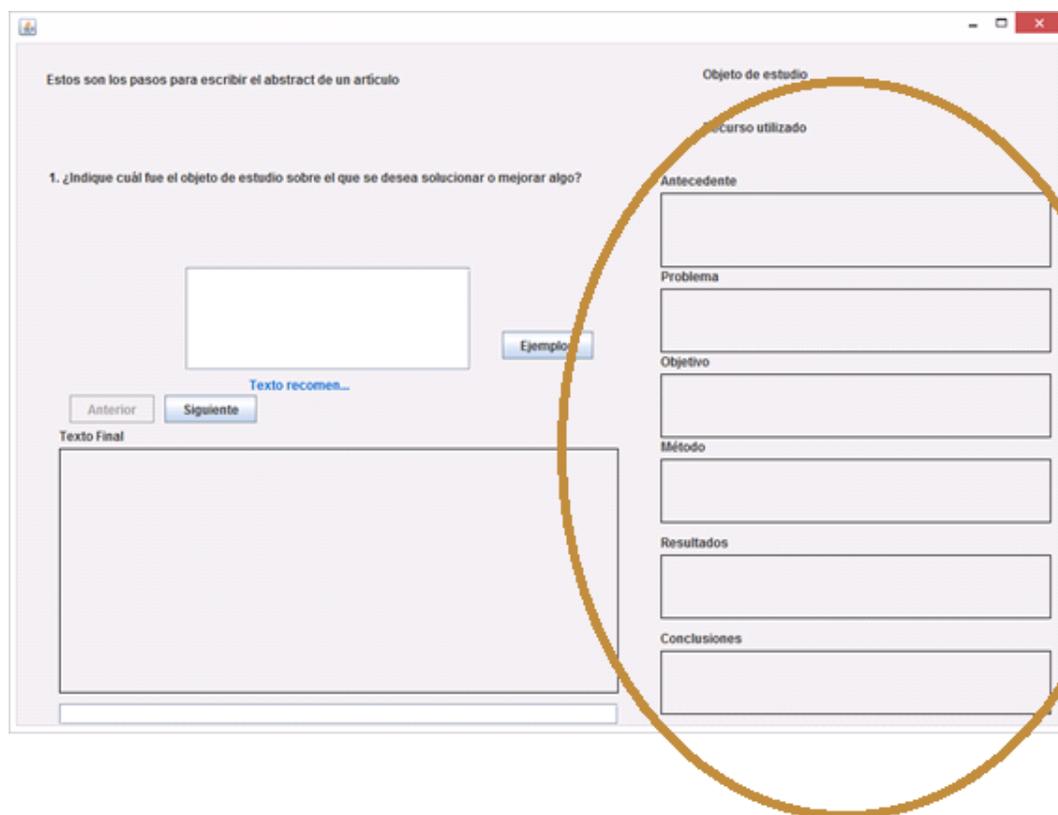
Texto Final

**Figura 13 Preguntas generadas**

Como se describió en el diseño, el prototipo realiza las tareas de la macroplanificación, para lo cual se reproduce el modelo que se ha diseñado. El modelo determina el orden de las secciones encontradas, además de los conectores entre secciones como se ve en la figura 9.

El prototipo solicita las meta-entidades descritas en el modelo. Con esta información se plantean las preguntas necesarias para la escritura del antecedente, problema, objetivo, entre otros. Con estas preguntas se intenta reducir la dificultad del escritor en el planteamiento de las oraciones, consiguiendo así, realizar la fase de microplanificación.

Cada sección se describe en el panel derecho con el objeto de diferenciar cada una durante el proceso de escritura (ver figura 14).



Estos son los pasos para escribir el abstract de un artículo

1. ¿Indique cuál fue el objeto de estudio sobre el que se desea solucionar o mejorar algo?

Objeto de estudio

Recurso utilizado

Antecedente

Problema

Objetivo

Método

Resultados

Conclusiones

Anterior

Siguiente

Texto Final

Texto recomen...

Ejemplo

**Figura 14 Panel de identificación de secciones**

De esta manera se puede identificar cada sección cuando sea necesario sin tener que leer todo el texto. Como en el caso de las conclusiones que se pide que relate las mismas tomando en cuenta el cumplimiento del objetivo, basándose en lo obtenido en los resultados.

Algunas de las características que plantea el prototipo es la sugerencia del tamaño de cada sección basado en el número promedio de caracteres encontrado en los artículos revisados (ver figura 15).

Estos son los pasos para escribir el abstract de un artículo

1. ¿Indique cuál fue el objeto de estudio sobre el que se desea solucionar o mejorar algo?

Objeto de estudio

Recurso utilizado

Antecedente

Problema

Objetivo

Método

Resultados

Conclusiones

Anterior Siguiente

Texto Final

Texto recomen...

Ejemplos

**Figura 15 Mensaje de número límite de caracteres de escritura**

El botón Ejemplos presenta un modelo de respuesta para cada pregunta para que el usuario tenga una referencia del tipo de oraciones que comúnmente se escribe. Finalmente, en el cuadro inferior se muestra todo el texto fusionado según las pautas del modelo (ver figura 16).

Anterior Siguiente

Texto Final

La falta de calidad en el producto software que desarrollan las empresas Sin embargo. Los sistemas de software presentan muchos defectos por la falta de aplicación de técnicas de desarrollo adecuadas Por lo tanto. Se realiza un experimento aplicando el desarrollo basado en pruebas TDD, y se compara con técnicas tradicionales para mediar la calidad externa del código y productividad de los desarrolladores Para lo cual. Se realizó un experimento mediante el desarrollo de programas y al final se evaluaron las variables respuesta calidad y productividad No se ha podido determinar si las variables calidad y productividad, inciden positiva o negativamente en la productividad y en la calidad externa Finalmente.

**Figura 16 Texto resultante del proceso de escritura**

## CAPÍTULO V

### VALIDACIÓN DEL MODELO

Para la validación del modelo se realizó un ejercicio donde investigadores experimentados hicieron pruebas con el prototipo y luego del resultado obtenido manifestaron sugerencias de mejora que sirvieron como retroalimentación para el mismo. De igual forma, el prototipo fue utilizado por estudiantes de investigación que también brindaron su criterio y sugerencias para que el prototipo sea fácilmente adaptable para investigadores que no cuenten con experiencia. De esta manera, se encontraron algunas limitaciones en la propuesta, ya que en el caso de los investigadores con experiencia solicitaban mayor cantidad de conectores entre las oraciones, para que de esta forma ellos puedan escoger cuál de ellos conecta mejor las oraciones.

También se encontró que algunas preguntas fueron ambiguas por lo que hubo que reformular algunos términos, y acotarlos mejor. Con la retroalimentación de los investigadores también se aumentaron los ejemplos descritos en el software, que en principio no incluía el modelo.

Para los investigadores experimentados la herramienta no fue de gran utilidad debido a su experiencia escribiendo. Sin embargo resaltan la utilidad del modelo al momento de escoger el orden de las secciones, y las relaciones abstractas entre las secciones, como contraste o elaboración que son conceptos de la Teoría de la Estructuración Retórica.

Para los estudiantes de investigación que tienen poca experiencia en la escritura de artículos científicos la herramienta fue de gran ayuda, ya que al tener preguntas específicas y puntuales les facilitó en gran medida escribir acerca de sus investigaciones.

De esta manera la herramienta muestra eficiencia al momento de realizar la tarea de macroplanificación y no así la tarea de microplanificación para los investigadores experimentados.

## 5.1. Pruebas

### Sujeto de Pruebas 1: Investigador Experimentado

Estos son los pasos para escribir el resumen de un artículo

10 Escriba las conclusiones en función del cumplimiento del objetivo basado en los resultados.

Ejemplos

Texto recomendado: 30

Anterior Siguiente

Texto Final

quebres de señal inalámbrica por cobertura que afectan al servicio de videostreaming. Sin embargo. Los factores externos que afectan al canal inalámbrico como ruido, interferencia, cobertura, entre otros, afectan a la experiencia de usuario en un servicio de videostreaming debido a que tienen que visualizar nuevamente todo el contenido. Por lo tanto. El objetivo de la propuesta es proveer de una solución que permita ser proactivo ante un posible quiebre del canal inalámbrico para evitar pérdidas de información y si es el caso de un quiebre inminente almacenar la información en origen hasta reestablecer la sesión y transmitir la información almacenada y continuar con el resto. Para lo cual. Se utilizó la metodología design science para implementar un prototipo que permita recrear los escenarios a ser evaluados Dentro de los resultados obtenidos se destaca la mejora en los tiempos de reconexión de las sesiones de videostreaming y la mínima pérdida de datos al cliente, generando una calidad de experiencia en usuario. Finalmente. Con la solución propuesta se mitigó las interrupciones de la señal wifi por cobertura dado la proactividad que el sistema presenta al cliente mediante alertas que permiten mantenemos en zonas de cobertura

presenta al cliente mediante alertas que permiten mantenemos en zonas de cobertura-htm

Objeto de estudio

Disrupciones en redes inalámbricas para servicio de videostreaming

Recurso utilizado

modelo basado en webrtc y realidad aumentada

Antecedente

quebres de señal inalámbrica por cobertura que afectan al servicio de videostreaming

Problema

Los factores externos que afectan al canal inalámbrico como ruido, interferencia, cobertura, entre otros,

Objetivo

El objetivo de la propuesta es proveer de una solución que permita ser proactivo ante un posible quiebr

Método

Se utilizó la metodología design science para implementar un prototipo que permita recrear los escene

Resultados

Dentro de los resultados obtenidos se destaca la mejora en los tiempos de reconexión de las sesione

Conclusiones

Con la solución propuesta se mitigó las interrupciones de la señal wifi por cobertura dado la proactivi

**Figura 17 Test de validación. Sujeto de pruebas 1**

Quebres de señal inalámbrica por cobertura que afectan al servicio de videostreaming. Sin embargo. Los factores externos que afectan al canal inalámbrico como ruido, interferencia, cobertura, entre otros, afectan a la experiencia de usuario en un servicio de videostreaming debido a que tienen que visualizar nuevamente todo el contenido. Por lo tanto. El objetivo de la propuesta es proveer de una solución que permita ser proactivo ante un posible quiebre del canal inalámbrico para evitar pérdidas de información y si es el caso de un quiebre inminente almacenar la información en origen hasta reestablecer la sesión y transmitir la información almacenada y continuar con el resto. Para lo cual. Se utilizó la metodología design science para implementar un prototipo que permita recrear los escenarios a ser evaluados Dentro de los resultados obtenidos se destaca la mejora en los tiempos de reconexión de las sesiones de videostreaming y la mínima pérdida de datos al cliente, generando una calidad de experiencia en usuario. Finalmente. Con la solución propuesta se mitigó las interrupciones de la señal wifi por cobertura dado la proactividad que el sistema presenta al cliente mediante alertas que permiten mantenemos en zonas de cobertura

de experiencia en usuario. Finalmente. Con la solución propuesta se mitigó las interrupciones de la señal wifi por cobertura dado la proactividad que el sistema presenta al cliente mediante alertas que permiten mantenernos en zonas de cobertura.

## Observaciones del Sujeto de Pruebas 1

- La herramienta es bastante útil, ya que permite realizar un resumen claro y preciso. Las preguntas generadas en el prototipo son precisas y permiten dar respuestas concretas.
- Se recomienda utilizar más conectores para las oraciones.
- Se recomienda empezar por el problema para dar respuesta con mayor facilidad a las demás preguntas.

## Sujeto de Pruebas 2: Estudiante de Investigación

Estos son los pasos para escribir el abstract de un artículo

10. Escriba las conclusiones en función del cumplimiento del objetivo basado en los resultados.

Texto recomendado: 234

Anterior Siguiente Ejemplos

Objeto de estudio  
Portal web MIESPE  
Recurso utilizado  
Heurísticas y pruebas de usabilidad

Antecedente

La usabilidad de un sistema es un factor determinante en el éxito o fracaso del mismo.

Problema

El portal web MIESPE al ser un sistema antiguo carece de buenas prácticas de usabilidad, lo cual ocasiona insatisfacción al momento en que los estudiantes de la ESPE lo utilizan para realizar algún tipo de actividad académica.

Objetivo

El presente trabajo tiene como objetivo evaluar la usabilidad del portal web MIESPE con el fin de identificar los problemas existentes y solventar los mismos mediante el desarrollo de una aplicación móvil multiplataforma utilizando un Diseño Centrado en el Usuario.

Método

Se utilizaron los lineamientos de usabilidad que en conjunto con los resultados obtenidos de la evaluación de usabilidad asegurarán una mejora en: (i) la eficiencia, (ii) eficacia y (iii) satisfacción al realizar las actividades.

Resultados

Posterior a una prueba de usabilidad con las mismas características de la evaluación al portal web, se determinó que efectivamente los usuarios se sintieron más cómodos al utilizar la aplicación móvil.

Conclusiones

La aplicación se puede solventar la problemática de usabilidad existente en el portal web causando satisfacción en los estudiantes al ser las aplicaciones móviles los dispositivos con

<html>La usabilidad de un sistema es un factor determinante en el éxito o fracaso del mismo.</html>

Figura 18 Test de validación. Sujeto de pruebas 2

La usabilidad de un sistema es un factor determinante en el éxito o fracaso del mismo. Sin embargo. El portal web MiESPE al ser un sistema antiguo carece de buenas prácticas de usabilidad, lo cual ocasiona insatisfacción al momento en que los estudiantes de la ESPE lo utilizan para realizar algún tipo de actividad académica. Por lo tanto. Se propone realizar una evaluación de usabilidad al portal web MiESPE con el fin de identificar los problemas existentes y solventar los mismos mediante el desarrollo de una aplicación móvil multiplataforma utilizando un Diseño Centrado en el Usuario. Para lo cual. Se requiere incluir en el diseño y desarrollo del mismo lineamientos de usabilidad que en conjunto con los resultados obtenidos de la evaluación de usabilidad asegurarán una mejora en: (i) la eficiencia, (ii) eficacia y (iii) satisfacción al realizar las actividades. Posterior a una prueba de usabilidad con las mismas características de la evaluación al portal web, se determinó que efectivamente los usuarios se sintieron más cómodos al utilizar la aplicación móvil. Finalmente. Al mantener un diseño centrado en el usuario en el desarrollo de la aplicación se puede solventar la problemática de usabilidad existente en el portal web causando satisfacción en los estudiantes al ser las aplicaciones móviles los dispositivos con los que más actividades realizan en la actualidad.

### **Observaciones del Sujeto de Pruebas 2.**

- Me pareció de gran utilidad pues orienta al usuario al permitirle centrarse en los objetivos planteados, y así describir de mejor manera lo que se desea expresar.
- Considerar los botones siguiente y atrás para mantener el texto que se escribió.
- Es mandatorio la longitud de caracteres descrita.

### Sujeto de Pruebas 3: Estudiante de Investigación

Estos son los pasos para escribir el abstract de un artículo

10. Escriba las conclusiones en función del cumplimiento del objetivo basado en los resultados.

Objeto de estudio  
alumnos de la Universidad de las Fuerzas Armadas ESPE

Recurso utilizado  
CRISP-DM

Antecedente  
LA EVALUACION QUE REALIZA EL CEAACES A INSTITUCIOONES DE EDUCACION SUPERIOR DE ACUERDO A LA TASA DE RETENCION GRADO Y TASA DE TITULACION GRADO CUMPLAN CON LOS PORCENTAJES ESTABLECIDOS

Problema  
LA UFA-ESPE ACTUALMENTE TIENE UN ALTO PORCENTAJE DE ALUMNOS DESERTORES Y POR ENDE SE TIENE PERDIDA DE PRESUPUESTO INVERTIDO EN LOS MISMOS

Objetivo  
SE IDENTIFICARAN LAS CAUSAN QUE INFLUYEN EN LA DESERCIÓN UNIVERSITARIA

Método  
se siguieron las seis etapas no rígidas junto con las tareas específicas que brinda la metodología

Resultados  
obteniendo un modelo cada departamento al que pertenece una carrera y otro para la desercion de la universidad que logran predecir con un bajo grado de error

Conclusiones  
se lograron identificar las principales causas de influencia y a la vez obtener un panorama claro de que esta pasando en la ESPE

Texto Final

LA EVALUACION QUE REALIZA EL CEAACES A INSTITUCIOONES DE EDUCACION SUPERIOR DE ACUERDO A LA TASA DE RETENCION GRADO Y TASA DE TITULACION GRADO CUMPLAN CON LOS PORCENTAJES ESTABLECIDOS Sin embargo. LA UFA-ESPE ACTUALMENTE TIENE UN ALTO PORCENTAJE DE ALUMNOS DESERTORES Y POR ENDE SE TIENE PERDIDA DE PRESUPUESTO INVERTIDO EN LOS MISMOS Por lo tanto. SE IDENTIFICARAN LAS CAUSAN QUE INFLUYEN EN LA DESERCIÓN UNIVERSITARIA Para lo cual. se siguieron las seis etapas no rígidas junto con las tareas específicas que brinda la metodología obteniendo un modelo cada departamento al que pertenece una carrera y otro para la desercion de la universidad que logran predecir con un bajo grado de error Finalmente. se lograron identificar las principales causas de influencia y a la vez obtener un panorama claro de que esta pasando en la ESPE

Anterior

Texto recomendado: 234

Siguiente

Ejemplos

san de influencia y a la vez obtener un panorama claro de que esta pasando en la ESPE</html>

**Figura 19 Test de validación. Sujeto de pruebas 3**

La evaluación que realiza el CEAACES a instituciones de educación superior de acuerdo a la tasa de retención grado y tasa de titulación grado que cumplan con los porcentajes establecidos. Sin embargo. La Universidad de las Fuerzas Armadas - ESPE actualmente tiene un alto porcentaje de alumnos desertores y por ende se tiene pérdida de presupuesto invertido en los mismos. Por lo tanto. Se identificarán las causas que influyen en la deserción universitaria. Para lo cual, se siguieron las seis etapas no rígidas junto con las tareas específicas que brinda la metodología obteniendo un modelo de cada departamento al que pertenece una carrera y otro para la deserción de la universidad que logran predecir con un bajo grado de error. Finalmente, se lograron identificar las principales causas de influencia y a la vez obtener un panorama claro de que está pasando en la ESPE.

#### Observaciones del Sujeto de Pruebas 3.

- Mejorar la interfaz y al momento de especificar si se eligió una metodología o herramienta es conveniente que se permita seleccionar en lugar de escribir.

## Sujeto de pruebas 4: Estudiante de Investigación

Estos son los pasos para escribir el abstract de un artículo

10. Escriba las conclusiones en función del cumplimiento del objetivo basado en los resultados.

Texto recomendado: 141

Anterior Siguiete

Texto Final

los niños aprenden mejor cuando escuchan y observan objetos Sin embargo, el uso de las herramientas virtuales mejoran el aprendizaje de los niños de preescolar Por lo tanto, realizar una aplicación de realidad virtual Para lo cual, se utilizó unity para realizar la aplicación de VR, con objetos y sonidos respectivos para cada vocal mediante pruebas en el Liceo Naval a los niños de 5 y 6 años Finalmente, se concluyó que los niños memorizan los objetos que ven y los relacionan con una vocal o color

Objeto de estudio  
Aprendizaje de numeros y vocales mediante VR  
Recurso utilizado  
Unity

Antecedente  
los niños aprenden mejor cuando escuchan y observan objetos

Problema  
el uso de las herramientas virtuales mejoran el aprendizaje de los niños de preescolar

Objetivo  
realizar una aplicación de realidad virtual

Método  
se utilizó unity para realizar la aplicación de VR, con objetos y sonidos respectivos para cada vocal

Resultados  
mediante pruebas en el Liceo Naval a los niños de 5 y 6 años

Conclusiones  
se concluyó que los niños memorizan los objetos que ven y los relacionan con una vocal o color

[Luyo que los niños memorizan los objetos que ven y los relacionan con una vocal o color</html>](#)

**Figura 20 Test de validación. Sujeto de pruebas 4**

Los niños aprenden mejor cuando escuchan y observan objetos. Sin embargo, el uso de las herramientas virtuales mejora el aprendizaje en niños de preescolar. Por lo tanto, realizar una aplicación de realidad virtual. Para lo cual. Se utilizó Unity para realizar la aplicación de VR, con objetos y sonidos respectivos para cada vocal mediante pruebas en el Liceo Naval a los niños de 5 y 6 años. Finalmente. Se concluyó que los niños memorizan los objetos que ven y los relacionan con una vocal o color.

### Observaciones del Sujeto de Pruebas 4.

- La herramienta es innovadora y útil si se tiene conocimiento sobre investigación, de lo contrario resulta en un proceso complejo.

## CAPÍTULO VI

### CONCLUSIONES Y FUTURAS LÍNEAS DE INVESTIGACIÓN

#### 6.1. Conclusiones

Las conclusiones manifiestan el resultado del proceso de investigación ejecutado, detallando la consecución de los objetivos específicos y de manera intrínseca el del objetivo general. Es importante el poder precisar que del estudio se concluye lo siguiente:

Conforme a los procesos de la Generación del lenguaje natural, es posible proponer una forma de escribir el resumen de un artículo científico teniendo en cuenta el orden de las secciones y las meta-entidades, ya que debido a la naturaleza y fuerte estructura exigida en los artículos científicos, es viable establecer y encontrar patrones con un gran número de coincidencia; uno de los patrones encontrados fue el orden de las secciones en el resumen, siendo el mismo: Antecedente, problema, objetivo, metodología, resultados y conclusiones. Otro patrón que se distinguió con mucha frecuencia fue el de las metaentidades como es el caso de un artefacto o recurso usado para la solución de problemas, por ejemplo, técnicas del procesamiento del lenguaje natural o preguntas automáticas adaptadas al contexto. Además, se encontraron meta-entidades que representan el objeto de estudio al que se desea modificar en su contexto, por ejemplo, coherencia en los textos de niños de primaria o la comunicación científica.

Mediante la teoría de la estructuración retórica es posible establecer un patrón que consista en la forma en la que se conectan las secciones del resumen obteniendo un flujo coherente y formalizado que denote cómo enlazar las ideas de tal forma que la lectura sea clara y precisa permitiendo demostrar y aceptar la primera parte de la hipótesis planteada.

Se ha comprobado que el diálogo puede comportarse como una herramienta de extracción de información, en el presente estudio fue posible desarrollar un dialogo basado en la estructura de un árbol de preguntas que fue diseñado conforme a los patrones lingüísticos encontrados en el cual los investigadores fueron la fuente de información del que el diálogo se alimenta permitiendo demostrar y aceptar la segunda parte de la hipótesis H1 (Sección ).

En base a la información extraída por el diálogo es posible generar un texto claro, sencillo y ordenado que está compuesto por ideas correctamente enlazadas y que permiten al lector un entendimiento general del estudio que un investigador pretenda comunicar permitiendo demostrar y aceptar la hipótesis.

Finalmente, la validación del modelo a través del prototipo que fue analizado y corroborado por investigadores experimentados y por estudiantes de investigación permitió comprobar que los procesos de la generación del lenguaje natural y los sistemas de tutoría inteligente combinados, facilitan al escritor del artículo científico enfocarse en lo que quiere comunicar, sin tener la necesidad de concentrarse en la determinación del contenido y la estructuración del discurso.

## **6.2. Futuras Líneas de Investigación**

Como futuras líneas de investigación, se propone el estudio de más patrones en los que se tome en cuenta el entendimiento de sus lectores, por ejemplo, el orden de las secciones ayudan al entendimiento de los lectores debido a que en el flujo de la escritura, el hecho de que una sección anteceda a otra, aporta información a la siguiente sección, de esta manera, se facilita la lectura y con ello el entendimiento. Así mismo, se debe tener en cuenta la facilidad de la lectura, para no caer en el aburrimiento de los lectores.

Se propone que se realice un estudio con un mayor número de artículos, y con ello crear un algoritmo o una red neuronal que aprenda como determinar las secciones del resumen de un artículo con lo cual se podrá generar un nuevo modelo basado en un análisis con un mayor número de ejemplos. Además, se sugiere determinar las diferencias entre los patrones lingüísticos usados en estudios de inglés y español.

Así como el diálogo es la forma más natural para la comunicación, la generación de lenguaje natural, es la forma más precisa de realizar escritos que cumplan con los requisitos de la comunicación científica, por lo cual, se recomienda realizar más estudios de la comunicación científica, consiguiendo así, diseñar un generador de lenguaje natural en el dominio de la comunicación científica.

Finalmente, cabe recalcar la utilidad que tiene el crear un sistema en el que interactúen varias arquitecturas, en este caso fue la interacción de un sistema de generación del lenguaje natural con un diálogo generado por un sistema de tutoría inteligente. Para trabajos futuros se propone realizar estudios en mayor profundidad para mejorar la interacción entre estos sistemas, con el objetivo de generar escritos semiautomáticos, basados en la información obtenida de personas, al estilo de un articulista.

## REFERENCIAS BIBLIOGRÁFICAS

- Al Emran, M., & Shaalan, K. (2014). A Survey of Intelligent Language Tutoring Systems. *IEEE: International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 393-399. Recuperado el Enero 4, 2017
- Álvarez Munárriz, L. (1994). *Fundamentos de Inteligencia Artificial*. Murcia, España. Recuperado el Julio 20, 2017
- Amador Hidalgo, L. (1996). *Inteligencia artificial y sistemas expertos*. Córdoba. Recuperado el Julio 26, 2017
- Benz, C. R., & Newman, I. (2008). Mixed methods research: Exploring the interactive continuum. *SIU Press*.
- Biran, O., & McKeown, K. (2015, Septiembre). Discourse Planning with an N-gram Model of Relations. *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, 1973-1977. Recuperado el Enero 16, 2017
- Bolaños, B. (1996). *Comunicación Escrita* (Primera ed.). San José, Costa Rica. Recuperado el Julio 25, 2017
- Cargill, M., & Smernik, R. (2015, Septiembre 25). Embedding publication skills in science research training: a writing group programme based on applied linguistics framework and facilitated by a scientist. *HERDSA*. Recuperado el Enero 4, 2017
- Cataldi, Z., & Lage, F. J. (2010). Modelado del Estudiante en Sistemas Tutores Inteligentes. *TE & ET*. Recuperado el Julio 10, 2017
- Chowdhury, G. G. (2003). Natural language processing. *Annual Review of INFORMATION SCIENCE AND TECHNOLOGY*, 37(1), 51-89. doi:10.1002/aris.1440370103

- Day, R. A. (2005). *Como escribir y publicar trabajos científicos* (3era ed.). Recuperado el Julio 28, 2017
- Denning, P. J. (1997). A New Social Contract for Research. *Communications of the ACM*, 40(2), 132-134. Recuperado el Marzo 20, 2017
- Dik, S. C., Valverde, F., & Mingorance, L. M. (1981). *Gramática funcional*. Sociedad General Española de Librería. Recuperado el Julio 25, 2017
- Dimitromanolaki, A., & Androutsopoulos, I. (2003). Learning to Order Facts for Discourse Planning in Natural Language Generation. *Workshop on Natural Language Generation*. Recuperado el Enero 10, 2017
- Dodigovic, M. (2013). Intelligent Sentence Writing Tutor: A System Development Cycle. *International Journal of Artificial Intelligence in Education*, 22(3), 141-160. Recuperado el Enero 15, 2017
- Geithner, C., & Pollastro, A. (2015). Construction engaged learning in Scientific Writing. *Journal of Applied Research in Higher Education*. *Emerald Insight.*, 7(2), 292-307. doi:10.1108/JARHE-04-2014-0053
- Gemayel, R. (2016). How to write a scientific paper. *FEBS PRESS*, 3882-3885. Recuperado el Enero 4, 2017
- Gorrostieta, J., & López-López, A. (2016). Argumentation Identification for Academic Support in Undergraduate Writings. *In European Conference on Technology Enhanced Learning*, 98-109. Recuperado el Enero 23, 2017
- Harbusch, K., Itsova, G., Koch, U., & Kühner, C. (2008, Octubre). The Sentence Fairy: a natural-language generation system to support children`s essay writing. *Computer Assisted Language Learning*, 21(4), 339-352. doi:10.1080/09588220802343496
- Hastings, P., Hughes, S., Britt, A., Blaum, D., & Wallace, P. (2014, Junio). Toward automatic inference of causal structure in student essays. *International*

*Conference on Intelligent Tutoring Systems*, 266-271. Recuperado el Enero 10, 2017

Haugeland, J. (2003). *La Inteligencia Artificial* (Vol. IV). México: Siglo XXI Editores. Recuperado el Julio 20, 2017

Hendel, R. J. (2014). Enhancing Writing through Strengthened Executive Function. *Systemics, Cybernetics and Informatics*, 17-22. doi:1690-4524

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004, Marzo). Design Science in Information Systems Research. *MIS Quarterly*, 28(1), 75-105. Recuperado el Marzo 20, 2017

Jayalakshmi, N., Suvarna, K., Vidya, H., & Sujatha, C. (2015). Enhancing Research Proficiency in Postgraduate Engineering Students. *IEEE*. Recuperado el Enero 4, 2017

Kirsch, G., & Sullivan, P. A. (1992). Methods and methodology in composition research. *SIU Press*.

Kirschner, C., Eckle-Kohler, J., & Gurevych, I. (2015, Junio 04). Linking the Thoughts: Analysis of Argumentation Structures in Scientific Publications. *Association for Computational Linguistics: Proceedings of the 2nd Workshop on Argumentation Mining*, 1-11. Recuperado el Enero 05, 2017

Krause, S., Hennig, L., Moro, A., Weissenborn, D., Xu, F., Uszkoreit, H., & Navigli, R. (2016). Sar-graphs: A language resource connecting linguistic knowledge with semantic relations from knowledge graphs. *Web Semantics: Science, Services and Agents on the World Wide Web*, 37-38, 112-131. doi:10.1016/j.websem.2016.03.004

Le, N.-T., & Huse, N. (2016). Evaluation of the Formal Models for the Socratic Method. *Intelligent Tutoring Systems*, 69-78. doi:10.1007/978-3-319-39583-8\_7

- Le, N.-T., Nguyen, N.-P., Seta, K., & Pinkwart, N. (2014). Automatic question generation for supporting argumentation. *Vietnam J Comput Sci*. doi:10.1007/s40595-014-0014-9
- Leyva, S. (2013). El proceso de escritura y publicación de un artículo científico. *Revista Electrónica Educare.*, 17(1), 5-27. Recuperado el Enero 4, 2017
- Liu, C.-L., Lee, C.-H., & Ding, B.-Y. (2012). Intelligent computer assisted blog writing system. *Expert Systems with Applications*, 39(4), 4496-4504. doi:10.1016/j.eswa.2011.09.139
- Liu, C.-L., Lee, C.-H., Yu, S.-H., & Chen, C.-W. (2010). Computer assisted writing system. *Expert Systems with Applications*, 804-811. doi:10.1016/j.eswa.2010.07.038
- Liu, M., Calvo, R. A., & Rus, V. (2012). G-Asks: An Intelligent Automatic Question Generation System for Academic Writing Support. *Dialogue and Discourse*, 101-124. doi:0.5087/dad.2012.205
- Liu, M., Calvo, R. A., & Rus, V. (2014). Automatic Generation and Ranking of Questions for Critical Review. *Educational Technology & Society*, 333-346. doi:1176-3647
- Liu, M., Calvo, R., Aditomo, A., & Pizzato, L. A. (2012). Using Wikipedia and conceptual graph structures to generate questions for academic writing support. *IEEE: Transactions on Learning Technologies.*, 5(3), 251-263. doi:10.1109/TLT.2012.5
- Lochbaum, K. E. (1998). A Collaborative Planning Model of Intentional Structure. *Computational Linguistics*, 24(4), 525-572. Recuperado el Enero 10, 2017
- Loyola Illescas, E., Jara Cobos, R. V., Narváez, A. M., Romero Ortega, A., Padilla Verdugo, J., Farfán Pacheco, P., & Flores Ortíz, Á. C. (2014). *Miradas desde la educación superior en Ecuador* (Primera ed.). Quito, Ecuador: Editorial Universitaria Abya-Yala. Recuperado el Noviembre 15, 2016

- Mack, C. (2014). How to Write a Good Scientific Paper: Structure and Organization. *Micro/Nanolithography, MEMS, and MOEMS*, 13(4). Recuperado el Enero 4, 2017
- Mantilla-Villareal, A., Medina Fontalvo, J., Velasco-Bayuelo, C., Algarín Gregory, J., Rodelo Salcedo, E., de la Rosa Barranco, D., & Caballero-Uribe, C. (2010). Guía práctica para publicar un artículo en revistas latinoamericanas. *Salud Uninorte*, 5-27. Recuperado el Enero 4, 2017
- Mazidi, K., & Tarau, P. (2016). Automatic Question Generation: From NLU to NLG. *Intelligent Tutoring Systems*, 23-33. doi:10.1007/978-3-319-39583-8
- McNamara, D. S., Raine, R., Roscoe, R., Crossley, S. A., Jackson, T. G., Dai, J., . . . Graesser, A. C. (2012). The Writing-Pal: Natural Language Algorithms to Support Intelligent Tutoring on Writing Strategies. *Applied Natural Language Processing*, 298-311. doi:10.4018/978-1-60960-741-8.ch017
- Mu, J., van Aalst, J., Chan, C., & Fu, E. (2014, Junio). Automatic Coding of Questioning Patterns in Knowledge-building Discourse. *The 11th International Conference of Learning Sciences*, 23-27. Recuperado el Enero 20, 2017
- Nistor, N., Trăușan-Matu, Ș., Dascălu, M., Duttweiler, H., Chiru, C., Baltés, B., & Smeaton, G. (2015). Finding student-centered open learning environments on the internet: Automated dialogue assessment in academic virtual communities of practice. *Computers in Human Behavior*, 47, 119-127. doi:10.1016/j.chb.2014.07.029
- Norvig, P., & Russell, S. J. (2010). *Artificial Intelligence: A Modern Approach* (Tercera ed.). Prentice Hall. Recuperado el Julio 20, 2017

- Orwell's, G. (2007, Marzo 21). *Pick the Brain Grow Yourself*. Recuperado el Julio 2, 2017, de <https://www.pickthebrain.com/blog/george-orwells-5-rules-for-effective-writing/>
- Pabitha, P., Mohanna, M., Suganthi, S., & Sivanandhini, B. (2014). Automatic Question Generation System. *2014 International Conference on Recent Trends in Information Technology*. Recuperado el Enero 17, 2017
- Pollak, S., Lesjak, B., Kranjc, J., Podpecan, V., Nidaric, M., & Lavrac, N. (2015). RoboCHAIR: Creative Assistant for Question Generation and Ranking. *2015 IEEE Symposium Series on Computational Intelligence*, 1468-1475. doi:10.1109/ssci.2015.208
- Reiter, E., & Dale, R. (2000). *Building natural language generation systems*. Cambridge University Press. Recuperado el Enero 15, 2017
- Renkema, J. (1999). *Introducción a los estudios del discurso*. Recuperado el Julio 29, 2017
- Roscoe, R. D., Allen, L. K., Weston, J. L., Crossley, S. A., & McNamara, D. S. (2014). The Writing Pal Intelligent Tutoring System: Usability Testing and Development. *Computers and Composition*, 34, 39-59. doi:10.1016/j.compcom.2014.09.002
- Thirugnanasambantham, C. (2011). An Intelligent System Framework for an Automated Language Tutoring Tool. *IEEE: First International Conference on Informatics and Computational Intelligence*, 122-125. doi:10.1109/ICI.2011.29
- Trausan-Matu, S., Dascalu, M., & Dessus, P. (2012). Textual Complexity and Discourse Structure in Computer-Supported Collaborative Learning. *Springer-Verlag Berlin Heidelberg*, 352-357. Recuperado el Enero 05, 2017

- Universidad del Rosario. (2003). Como plantear preguntas de investigación. Recuperado el Enero 6, 2017, de [http://datateca.unad.edu.co/contenidos/100104/100104\\_EXE/leccin\\_11\\_la\\_pregunta\\_de\\_investigacin.html](http://datateca.unad.edu.co/contenidos/100104/100104_EXE/leccin_11_la_pregunta_de_investigacin.html)
- Van Dijk, T. A., & Mayoral, J. A. (1987). *Pragmática de la comunicación literaria*. Arco/Libros. Recuperado el Julio 10, 2017
- Vega Mora, H. (2008, Octubre 3). *Sintáctica, Semántic y Pragmática*. Recuperado el Julio 15, 2017, de Freemasonry: [http://www.freemasons-freemasonry.com/Vega\\_Mora.html](http://www.freemasons-freemasonry.com/Vega_Mora.html)
- Vicente, M., Barros, C., Peregrino, F. S., Agulló, F., & Lloret, E. (2015). La generación de lenguaje natural: análisis del estado actual. *Computación y Sistemas*, 19(4), 721-756. doi:10.13053/CyS-19-4-2196
- Wieringa, R. J. (2014). Design Science Methodology for Information Systems and Software Engineering. *Springer*. doi:10.1007/978-3-662-43839-8
- Wilcock, G., & Yamamoto, S. (2015). Towards computer-assisted language learning with robots, wikipedia and CogInfoCom. *2015 6th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. doi:10.1109/coginfocom.2015.7390575