



SEP
SECRETARÍA DE
EDUCACIÓN PÚBLICA



TECNOLÓGICO NACIONAL DE MÉXICO
Instituto Tecnológico de Hermosillo

Instituto Tecnológico de Hermosillo

División de Estudios de Posgrado e Investigación

Diseño e implementación de algoritmo para el análisis de la Instrumentación
Didáctica de asignaturas de educación superior tecnológica utilizando
ontologías y Procesamiento de Lenguaje Natural

TESIS

Presentada por:

Ing. Daniel Alfredo Hernández Carrasco

Como requisito parcial para
Obtener el grado de

Maestro en Ciencias de la Computación

Director de tesis: M.C. César Enrique Rose Gómez

Codirector de tesis: Dr. Samuel González López

Hermosillo, Sonora, México

Agosto de 2020



Av. Tecnológico S/N Col. El Sahuaro, C.P. 83170 Hermosillo, Sonora.

Tel. (662) 2-606500 Ext. 136

E-mail: depi_hermosillo@tecnm.mx





EDUCACIÓN

SECRETARÍA DE EDUCACIÓN PÚBLICA



TECNOLÓGICO NACIONAL DE MÉXICO

Instituto Tecnológico de Hermosillo

"2020, Año de Leona Vicario, Benemérita Madre de la Patria"

SECCIÓN: DIV. EST. POS. E INV.
No. OFICIO: DEPI/067/20
ASUNTO: AUTORIZACIÓN DE IMPRESIÓN DE TESIS.

25 de Junio de 2020

**C. DANIEL ALFREDO HERNÁNDEZ CARRASCO,
PRESENTE.**

Por este conducto, y en virtud de haber concluido la revisión del trabajo de tesis que lleva por nombre "ALGORITMO PARA EL ANÁLISIS DE LA INSTRUMENTACIÓN DIDÁCTICA DE ASIGNATURAS DE EDUCACIÓN SUPERIOR TECNOLÓGICA UTILIZANDO ONTOLOGÍAS Y PROCESAMIENTO DE LENGUAJE NATURAL", que presenta para el examen de grado de la MAESTRÍA EN CIENCIAS DE LA COMPUTACIÓN, y habiéndola encontrado satisfactoria, nos permitimos comunicarle que se autoriza la impresión del mismo a efecto de que proceda el trámite de obtención de grado.

Deseándole éxito en su vida profesional, quedo de usted.

ATENTAMENTE

M.C. CÉSAR ENRIQUE ROSE GÓMEZ
DIRECTOR

A. Mancinas

DR. ABELARDO MANCINAS GÓZALEZ
SECRETARIO

M.C. ANA LUISA MILLÁN CASTRO
VOCAL

M.C.O. ROSA IRENE SÁNCHEZ FERMÍN
JEFA DE LA DIVISION DE ESTUDIOS DE POSGRADO E INVESTIGACIÓN



S.E.P.

INSTITUTO TECNOLÓGICO
DE HERMOSILLO
DIVISIÓN DE ESTUDIOS
DE POSGRADO

RISF/fjcg*



Av. Tecnológico S/N Col. El Sahuaro C.P. 83170 Hermosillo, Sonora
Tel. 01 (662) 2-606500, Ext. 136 e-mail: depi_hermosillo@tecnm.mx
www.tecnm.mx | www.ith.mx



DEDICATORIA

Antes que nada, quisiera agradecer a mi madre, que me apoyo en la decisión de continuar mis estudios y me motivo con cariño para que creciera como persona, acompañándome a cada momento a pesar de la distancia. De igual forma, a mis ángeles que me cuidan desde el cielo, quienes me apoyaron y me desearon lo mejor y, aunque ahora no estén presentes para presenciar este logro, les dedico este fragmento como muestra de que nunca les olvidare.

A la par, le agradezco a mi compañera de vida Diana León, quien es un sustento importante en mi vida, quien a pesar de la distancia y los momentos de incertidumbre me ha brindado su amor y comprensión, motivándome a superarme día a día a pesar de las adversidades, a ver la vida de una forma diferente y a confiar en mis capacidades y decisiones. Por todas las veces que leíste mis borradores, por motivarme y darme esa compañía perfecta, te lo agradezco amor.

Tambien, quiero agradecer a mi director de tesis M. Cesar Rose, quien me guio en este proceso brindándome su conocimiento para hacer posible el proyecto, gracias por sus observaciones y tomarse el tiempo para aconsejarme, sus enseñanzas me acompañaran el resto de mi vida. De igual forma, a mi codirector Dr. Samuel González, quien creyó en mis capacidades y me motivó a seguirme preparándome y creciendo en el ámbito académico. De modo idéntico, al Dr. Aurelio López que me recibió en el INAOE como parte de la estancia que me permitió comprender nuevas metodologías que me ayudarían a alcanzar los resultados finales de este proyecto. A todos ustedes, gracias por ser los formadores y brindarme su apoyo y tiempo para concluir este proyecto.

Finalmente, a mi familia, amigos y compañeros con quienes compartí momentos que son mas que recuerdos, gracias por su apoyo para alcanzar las metas y la confianza en que haría un buen trabajo. Muchas gracias a todos mis seres queridos por estar presente en esta etapa importante en mi vida, por lo que les dedico de todo corazón el resultado de mi proyecto: mi tesis.

Ing. Daniel Alfredo Hernández Carrasco

Índice de contenido

Capítulo I	1
1.1- Introducción.....	1
1.2- Antecedentes.....	2
1.3- Planteamiento del Problema	4
1.4- Objetivos.....	7
1.4.1- Objetivo general	7
1.4.2- Objetivos específicos.....	7
1.5- Justificación de la tesis	7
1.6- Alcances y delimitaciones	8
1.7- Metodología.....	8
1.8- Organización de la tesis.....	9
Capítulo II-Marco Teórico	11
2.1- Introducción.....	11
2.2- Modelo Educativo en México.....	12
2.2.1- Modelo EBC.....	12
2.2.2- Puntos del Modelo ET y Modelo EBC.....	13
2.2.3- Principio pedagógico del Modelo EBC	14
2.2.4- Instrumentación Didáctica.....	15
2.2.5- Taxonomía de Bloom	17
2.3- Modelos de Conocimiento.....	18
2.3.1- Sistemas basados en conocimiento.....	18
2.3.2- Razonamiento	20
2.3.3- Ontologías.....	21
2.3.4- Componentes de una Ontología.....	22
2.3.5- Tipos de Ontologías.....	23

2.3.6- Metodologías para Ontologías.....	24
2.3.7- Lenguajes y herramientas para Ontologías.....	24
2.4- Procesamiento de Lenguaje Natural.....	26
2.4.1- Niveles de análisis y Normalización	28
2.4.2-Similitud entre cadenas.....	29
2.4.3- Librerías de PLN	30
2.4.4- Algoritmos de clasificación	32
Capítulo III- Análisis y diseño.....	34
3.1- Introducción.....	34
3.2- Arquitectura del sistema	35
3.3- Diseño de la base de datos.....	36
3.4- Diseño y procesos de la ontología.....	37
3.4.1- Metodología Methontology	37
3.4.2- Conceptualización	38
3.4.3- Procesos	43
3.5- Diseño de algoritmo de PLN	46
3.5.1- Autómata	48
3.5.2-Composición de una ID.....	51
3.5.3-Match de actividades de aprendizaje.....	52
3.5.4-Proceso de Match.....	52
3.5.3- Evidencia de actividades	53
3.5.4- Instrumentos de evaluación de evidencias.....	55
Capítulo IV-Implementación	56
4.1-Introducción.....	56
4.2- Implementación de la Ontología.....	57
4.2.1.- Construcción de las clases	57

4.2.2- Relaciones entre conceptos.....	58
4.2.3- Propiedades de datos	59
4.2.4- Instancias	59
4.2.5- Motor de inferencia	61
4.2.6- Visualización gráfica de la ontología	61
4.3- Algoritmo de PLN	62
4.3.1.-Requisitos previos.....	62
4.3.2- Autómata para segmentación de competencia.....	63
4.3.3- Consultar Ontología.....	65
4.3.4- Match	66
4.2.5.-Relación entre Actividades y Evidencias de Aprendizaje	68
4.2.6- Relación entre Evidencias e Instrumentos.....	70
Capítulo 5-Análisis de resultados	71
5.1.-ID-Prueba I	72
5.2.- Pruebas al autómata	72
5.3.-Pruebas Match	73
5.4.-Prueba de relación entre actividades y evidencias.....	76
5.5.-Prueba de evaluación de evidencias	76
5.6.-Resultados enviados al módulo de recomendación	77
5.7.-ID-Prueba piloto	80
5.8-Comparativa de resultados.....	86
5.9-Discusión de resultados	90
Capítulo 6-Conclusiones	93
6.1.-Conclusiones.....	93
6.2.-Trabajo a Futuro	95
Bibliografía	97
Anexos	104

Índice de Figuras

Figura 1 -Arquitectura completa del sistema.....	35
Figura 2 -Diagrama de la base de datos del sistema.....	37
Figura 3 -Diagrama de relaciones binarias	41
Figura 4 -Vista de la Taxonomía de Bloom en la ontología.....	42
Figura 5 -Diagrama de contexto (Nivel 0).....	43
Figura 6 -Diagrama de nivel inferior (Nivel 1).....	44
Figura 7 -Diagrama de nivel superior (Nivel 2)	45
Figura 8 -Secuencia de algoritmo de análisis de PLN	47
Figura 9 -Secuencia de análisis competencia específica	48
Figura 10 -Autómata para segmentar competencia específica	49
Figura 11 -Ejemplo de análisis Morfológico de la competencia específica	50
Figura 12 -Ejemplo de tabla que contiene los datos de una ID	51
Figura 13 -Ejemplo de matriz de evaluación de una ID	51
Figura 14 -Proceso de match de actividades de aprendizaje	52
Figura 15 -Revisión de redacción de las actividades utilizando Match.....	53
Figura 16 -Relación entre actividades y evidencias de aprendizaje	54
Figura 17 -Vista parcial de los conceptos de la ontología	58
Figura 18 -Vista parcial de las relaciones entre las clases.....	58
Figura 19 -Vista parcial de los atributos de las clases	59
Figura 20 -Vista parcial de las instancias	60
Figura 21 -Visualización de las clases y sus individuos.....	60
Figura 22 -Deducción obtenida utilizando el motor de inferencia	61
Figura 23 -Visualización de las relaciones utilizando OntoGraf.....	62
Figura 24 -Pseudocódigo base del autómata.....	64
Figura 25 -Resultado de análisis morfológico	65
Figura 26 -Función para consultas a la ontología.....	66
Figura 27 -Pseudocódigo base para obtener el Match	67
Figura 28 -Seudocodigo base del análisis de relaciones.....	69

Figura 29- Seudocodigo base de evidencias e instrumentos.....	70
Figura 30- Resultado análisis de autómeta	73
Figura 31- Actividades asociadas en base al dominio de la competencia	73
Figura 32- Match entre actividad de aprendizaje y competencia	74

Índice de tablas

Tabla 1- Comparativa del Modelo ET y el Modelo EBC	14
Tabla 2- Glosario de conceptos y su descripción (Anexo 1)	39
Tabla 3- Glosario de individuos y sus sinónimos (Anexo 2).....	39
Tabla 4- Propiedades de las clases (Anexo 4)	40
Tabla 5- Relación entre conceptos (Anexo 5)	41
Tabla 6- Relación entre Evidencias e instrumentos (ver Anexo 6)	55
Tabla 7- Instrumentación didáctica de prueba	72
Tabla 8- Match de actividades de aprendizaje	74
Tabla 9- Match de actividades de enseñanza	75
Tabla 10- Match de evidencias de aprendizaje	75
Tabla 11- Match de instrumentos de evaluación	75
Tabla 12- Evidencias a partir de actividades	76
Tabla 13- Actividades a partir de evidencias	76
Tabla 14- Instrumento de evaluación de evidencias	77
Tabla 15- Variables resultantes de competencia específica.....	77
Tabla 16- Variables resultantes entre competencia y actividades.....	78
Tabla 17- Variables resultantes del análisis del Match.....	78
Tabla 18- Variables resultantes de las actividades y evidencias.....	79
Tabla 19- Variables resultantes entre instrumentos y evidencias	79
Tabla 20- Instrumentación didáctica de prueba piloto.....	81
Tabla 21- Match entre competencias y actividades (Prueba piloto)	82
Tabla 22- Match actividades de aprendizaje (Prueba piloto).....	83
Tabla 23- Match de actividades de enseñanza (Prueba piloto).....	83
Tabla 24- Match de evidencias de aprendizaje (Prueba piloto)	84
Tabla 25- Match de instrumentos de evaluación (Prueba piloto)	84
Tabla 26- Evidencias asociadas a actividades de aprendizaje (Prueba piloto)	85
Tabla 27- Instrumentos asociados a evidencias de aprendizaje (Prueba piloto).....	85
Tabla 28- Comparativa de competencia específica.....	86

Tabla 29- Comparativa de relación entre competencia y actividades.....	87
Tabla 30- Comparativa en match de actividades de enseñanza	87
Tabla 31- Comparativa en match de actividades de aprendizaje	88
Tabla 32- Comparativa en match de evidencias de aprendizaje	88
Tabla 33- Comparativa en match de instrumentos de evaluación	89
Tabla 34- Comparativa de relación entre actividades y evidencias	89
Tabla 35- Comparativa de relación entre evidencias e instrumentos.....	90
Tabla 36- Resultados finales	91

Capítulo I

1.1- Introducción

A nivel mundial, la integración de las Tecnologías de la Información y Comunicación (TIC) a los procesos educativos ha sido considerada una necesidad para el desarrollo progresivo de los procesos de enseñanza-aprendizaje, esto encaminado a la mejora en la calidad educativa y al cumplimiento de las exigencias en las actuales generaciones [1].

A pesar de los esfuerzos, realizar la transición no ha sido fácil debido a que algunos docentes aún tienen problemáticas al intentar adaptar a las TIC en su metodología de trabajo y en la selección adecuada de las herramientas didácticas que tiene disponibles en base a la actividad a realizar. Uno de los principales motivos, es que prefieren mantener sus métodos tradicionales de enseñanza, desaprovechando las oportunidades que otorga el uso de las nuevas tecnologías, dedicándose solo a digitalizar la información sin llegar a desarrollar la competencia digital [2][3][4].

De lo anterior resulta que, una de las principales causas de este estancamiento transitivo es el carecer de un perfil pedagógico (por parte de los profesionistas) que ayude a afrontar las necesidades del proceso de enseñanza-aprendizaje [5], obstaculizando la adaptación de estrategias de acuerdo con el Modelo Educativo Basado en Competencias (EBC) sobre el cual se basa la educación en México [6].

En conjunto, dichas acciones son una problemática que causa un retroceso en vez de un avance hacia un cambio de paradigma. Para poder salir del estancamiento, es necesario que el docente realice una planeación de sus cursos mediante una instrumentación didáctica (ID), donde se planteen las actividades de enseñanza-aprendizaje en conjunto con las evidencias de aprendizaje que se esperan obtener como retroalimentación por parte del alumno y los instrumentos con los cuáles se pretenden evaluar [4][7][8], teniendo una visión clara sobre la metodología a seguir y no una relativa sobre el procedimiento para una redacción ideal [9].

Ahora bien, cada institución tiene un departamento encargado de la evaluación de las ID, pero la carga de trabajo impide una rápida retroalimentación al docente, quien tiene que esperar días (inclusive semanas) para recibir la aprobación o la lista de ajustes, lo que implica un nuevo proceso para cumplir con los requerimientos del departamento.

En base a esa problemática, el presente trabajo expone el diseño e implementación de un módulo para el análisis de la ID de asignaturas de educación superior tecnológica utilizando técnicas de Procesamiento de Lenguaje Natural (PLN) y modelos de conocimiento (también llamados ontologías). Como tal, el módulo forma parte de una plataforma dedicada a la capacitación y formación de docente, donde se busca que aprenda una metodología ideal para la elaboración de su ID.

1.2- Antecedentes

Con el tiempo, el Modelo Educativo Tradicional (ET) se ha ido modificando a las necesidades tanto de los alumnos como del profesor, adaptando las técnicas de enseñanza-aprendizaje que más se adecúe al perfil de egreso esperado, innovando al hacer uso de la tecnología en el proceso educativo [2][6].

Ante estos avances, el uso de tecnologías en la educación no es algo nuevo. En la década de 1970 con el desarrollo de técnicas de Inteligencia Artificial (IA), fue posible la implementación de sistemas tutores inteligentes con el objetivo de fungir como guía de aprendizaje a los estudiantes, emulando la conducta del maestro mediante algoritmos

computacionales. Desde entonces, esta rama ha desarrollado tutores cada vez más avanzados que son capaces de adaptar estrategias de aprendizaje de acuerdo con las fortalezas, deficiencias y necesidades de los alumnos para aprender algún tema en particular, abriendo la posibilidad a nuevas líneas de investigación sobre el tema con algoritmos cada vez más complejos [10].

Aunque su utilización en alumnos ha sido su principal objetivo, es viable preguntarse ¿Puede ser un docente capaz de aprender la metodología para la redacción de su ID utilizando un sistema para su capacitación y formación pedagógica?

En el año 2008 [11] fue aprobada la reforma educativa promovida por la Secretaría de Educación Pública (SEP) y por consiguiente un nuevo paradigma educativo que recibió el nombre de Modelo Educativo Basado en Competencias (EBC). Desde entonces, la capacitación de los maestros se convirtió en una necesidad que conllevó a la realización de cursos y diplomados con mayor frecuencia para mantenerles actualizados, con la finalidad de otorgar una educación de calidad manejando los planes estratégicos asentados en competencias [6].

Es importante aclarar que, a pesar de que la idea inicial era buena, el cambio hacia un nuevo paradigma educativo ha sido todo un desafío a causa de que no todos los docentes reciben sus capacitaciones con la misma frecuencia, puesto que en México se tienen algunas regiones de difícil acceso que son menos capacitadas que el resto, generando una brecha educativa palpable dependiendo de su geolocalización.

Ante esta problemática, en el año 2014 [12] en el estado de Jalisco se efectuó una prueba piloto con el uso de una plataforma educativa para la capacitación de maestros de educación indígena en zonas de difícil acceso. Vale la pena aclarar que los resultados de la capacitación fueron favorables y motivaron al profesor a aceptar recibir capacitaciones en línea, abriendo la posibilidad a volverlas masivas y así combatir la problemática de la brecha educativa en cuanto a la constancia de las capacitaciones.

Para tener una panorámica general del problema de la capacitación, estadísticas proporcionadas por la SEP en el documento “*Principales cifras del Sistema Educativo Nacional 2018-2019*” [13] indican que, para el ciclo escolar 2018-2019 se contaba con 414,408 maestros en el nivel de educación superior impartiendo clases en 5,535 instituciones. El total de alumnos para ese periodo era de 761,856 (417,747 mujeres y 344,109 hombres). En el caso del estado de Sonora, la cantidad de profesores laborando a nivel licenciatura universitaria y tecnológica era de 8,047 dispersados en 127 planteles con un total de 102,687 alumnos (50,305 hombres y 52,382 mujeres).

Con esos datos podemos determinar que, capacitar a más de 400,000 profesores activos en el país es todo un reto para la SEP. Como consecuencia de no hacerlo, si el docente carece de una perspectiva didáctica para aprovechar al máximo las herramientas tecnológicas de su entorno, tendrá pobres resultados en el proceso de enseñanza-aprendizaje y la transición hacia el Modelo EBC seguirá siendo una meta a largo plazo, por lo que la capacitación y formación de pedagogos es una necesidad para poder garantizar una educación de calidad.

Como trabajo previo, en el año 2007 [14] en la universidad de Antioquia en Colombia, se diseñó un algoritmo de PLN orientado a ayudar al docente a localizar los objetos de aprendizaje almacenados en un repositorio. Utilizando técnicas de análisis lingüístico, se buscaban aquellos objetos que permitieran evaluar la competencia en base a su objeto, considerando para ello la taxonomía de Bloom. Dicho proyecto, abrió la posibilidad a la búsqueda de patrones para ayudar a docentes a planear sus cursos utilizando técnicas de PLN y modelos de conocimiento para la recuperación de la información.

1.3- Planteamiento del Problema

Actualmente en México, para ser un docente del nivel de educación superior no es un requisito tener una formación pedagógica para laborar. Dado que la mayoría de los que trabajan en este nivel son profesionistas egresados de alguna licenciatura o posgrado acorde a la materia a impartir, algunos tienen dificultades (en especial en los primeros años en la labor) para redactar la ID de sus cursos al desconocer la metodología que se debe de seguir para su elaboración.

Como resultado de ese desconocimiento, la ID más que un apoyo para cumplir con las competencias iniciales se convierte en una forma confusa de abordar las temáticas de acuerdo con el plan inicial, desencadenado (generalmente) por la falta de dominio de las técnicas de enseñanza que se adquieren durante una formación pedagógica.

Con el fin de contrarrestar esta deficiencia, la SEP da capacitaciones constantes para contribuir al desarrollo de las capacidades y las prácticas en base al Modelo EBC, tratando de profesionalizarle orientando su perfil hacia la pedagogía. No obstante, un curso de un par de días no es suficiente comparado con los años de enseñanza de una carrera pedagógica. Ahora bien, las constantes modificaciones en las reformas educativas en México han causado conflicto en el entendimiento de los principales puntos que se tienen que cumplir con las bases del Modelo EBC, impidiendo el cumplimiento de lo establecido.

Si bien, la idea original era encaminar al alumno hacia un pensamiento crítico y creativo que le permitiera comprender la realidad de su entorno, la relativa interpretación del significado de una competencia ha ocasionado desconcierto al intentar dominar el enfoque sin contar con una formación para lograrlo. A la par, se encuentran aquellos que siguen implementando el Modelo ET al no lograr adaptarse a los cambios, llegando a considerar innecesario adoptar el nuevo paradigma educativo.

El panorama tampoco es alentador para los nuevos docentes, quienes a pesar del dominio de las TIC y haber vivido la transición del modelo ET al Modelo EBC, no tienen la experiencia frente al aula y la mayor parte no logra dominar las técnicas pedagógicas para tener una correcta planeación del curso.

Cierto es que dominar el tema de la materia a impartir no garantiza que se tengan buenas prácticas de enseñanza ni viceversa, por lo que es necesario tener un balance entre el conocimiento que se tiene y el método de enseñanza que se empleará para obtener un resultado favorable. De igual importancia, la formación de su perfil pedagógico es una necesidad que conlleva al desarrollo de habilidades que permitan enseñar en el aula aprovechando los recursos didácticos disponibles y así, poder transmitir de manera más eficiente el conocimiento hacia sus alumnos.

Dada esta problemática, el presente trabajo expone la construcción de un módulo de análisis que incluye técnicas de Procesamiento de Lenguaje Natural (PLN) y modelos de conocimiento para analizar y evaluar ID de asignaturas de educación superior tecnológica. El algoritmo formará parte de una plataforma donde el docente hará la redacción de su instrumentación que será analizada por el módulo. Posteriormente, el resultado obtenido será enviado a un módulo de recomendación (ajeno al de análisis), el cual se encargará de otorgar la retroalimentación para que se realicen los ajustes necesarios antes de su revisión final a cargo del departamento asignado por la institución.

La finalidad de la plataforma de capacitación y formación docente es que aprenda y comprenda la metodología para el desarrollo de una ID de acuerdo con el Modelo EBC, realizando una mejor estructuración de la planeación de sus cursos. A su vez, ayudará a dar una retroalimentación inmediata de la evaluación, algo que en la actualidad requiere de días para obtenerse por parte del departamento capacitado para la revisión al cual se le disminuirá la carga de trabajo, haciéndoles participe solo de la evaluación de la versión final del documento después de haber aprobado los filtros de la plataforma.

Hoy en día, es común el uso de plataformas educativas. Sin embargo, aplicado a la formación pedagógica es un área novedosa, por lo que esto nos lleva a las siguientes preguntas de investigación.

La pregunta principal que guía la investigación es:

¿Cuál es la configuración de un algoritmo para el análisis de la ID de asignaturas de educación superior tecnológica?

Preguntas secundarias:

- ¿Cómo diseñar un modelo ontológico que permita representar el conocimiento de una ID?
- ¿Cuál es el conocimiento normativo adicional a la ontología para realizar el razonamiento en el análisis de la ID?

- ¿Qué técnicas de PLN permiten encontrar patrones en el texto de cada una de las secciones que conforman la ID?

1.4- Objetivos

En este apartado del trabajo de tesis, se detallan tanto el objetivo general como los específicos planteados para su desarrollo.

1.4.1- Objetivo general

Diseñar e implementar un algoritmo para el análisis de la ID de asignaturas de educación superior tecnológica utilizando ontologías y Procesamiento de Lenguaje Natural.

1.4.2- Objetivos específicos

- Diseñar un modelo ontológico que permita representar el conocimiento de los elementos que conforman una ID.
- Identificar el conocimiento normativo adicional al modelo de conocimiento para realizar el razonamiento en el análisis de la ID.
- Definir las técnicas de PLN que permitan encontrar patrones en el texto de cada una de las secciones que conforman la ID.

1.5- Justificación de la tesis

El Modelo EBC propuesto por la SEP en el año 2008 para satisfacer las necesidades de los estudiantes, contempla al docente como un orientador y no el actor principal en un curso. En él, el profesor debe plantear competencias a cumplir en el proceso de enseñanza aprendizaje para que el estudiante desarrolle habilidades y aptitudes en la clase haciendo uso de una ID para planificar su curso.

Sin embargo, realizarla es todo un reto para aquellos que no tienen las bases pedagógicas ni la metodología a seguir, por lo que la elección incorrecta de las competencias y actividades de un curso son puntos sensibles por tomar en cuenta, dependiendo en gran medida el éxito o fracaso de un curso a la par con el aprovechamiento de los recursos didácticos y las habilidades pedagógicas para abordar las temáticas.

La finalidad del módulo a diseñar es analizar las ID en base a rúbricas obtenidas de evaluadores capacitados, conocimiento que se computarizará utilizando técnicas de PLN y modelos de conocimiento. Así, el proceso de evaluación se semi automatizará, reduciendo tanto la carga de trabajo como el personal dedicado a la revisión, dándole mayor fluidez al proceso.

Cabe mencionar que el módulo es un apoyo para los especialistas en el proceso de evaluación, por lo que no los sustituirá por completo, solo servirá como apoyo a los maestros para entregar la planeación final del curso de acuerdo con las rúbricas establecidas, pero son los especialistas quienes serán los encargados de otorgar la evaluación definitiva.

1.6- Alcances y delimitaciones

La presente investigación tiene el propósito de diseñar un algoritmo inteligente basado en técnicas de PLN y un modelo ontológico para la evaluación de la ID de asignaturas de educación superior tecnológica.

El proyecto está restringido a la aplicación del algoritmo basándose en rúbricas del personal certificado para evaluar ID redactadas por docentes de educación superior del área tecnológica, limitando el análisis a este dominio en específico.

1.7- Metodología

El presente trabajo, representa una investigación donde se analizan los contenidos de las secciones de la ID en base a rúbricas de evaluadores certificados para su revisión. Para lograrlo, el trabajo se desarrollará en cinco etapas:

Consulta bibliográfica: En la primera etapa se recopilará información del estado del arte del Modelo EBC, así como una breve comparativa con el Modelo ET. Posteriormente, se hará mención sobre lo referente a los modelos de conocimientos, sus componentes, tipos y algunas de las metodologías más comunes para su construcción. Finalmente, se documentarán técnicas de PLN para el análisis de textos en formato libre en el idioma español, incluyendo un breve apartado sobre algoritmos de clasificación.

Diseño y construcción del modelo ontológico y algoritmo: En la segunda etapa, se diseñará la arquitectura del sistema, así como la metodología para la construcción del modelo de conocimiento basado en los elementos de la ID y la definición de las reglas necesarias para inferir conocimiento. Después, se diseñará el algoritmo utilizando técnicas de PLN para la detección de patrones en las secciones de la ID.

Implementación del modelo Ontológico y algoritmo: En la tercera etapa se construirá la ontología en base a la metodología planteada en la etapa anterior, desde la definición de las clases hasta las relaciones entre los conceptos que le componen. También, se codificará el algoritmo de PLN para la detección de patrones en la competencia específica, las secciones de la ID y las relaciones entre dichas secciones.

Pruebas y análisis de resultados: Posteriormente, se utilizarán ID redactadas por docentes para probar las distintas funcionalidades del algoritmo, con el fin de evaluar el cumplimiento de los objetivos planteados.

Conclusiones: Para terminar, se analizarán los resultados obtenidos seguido de la redacción de las conclusiones y premisas de trabajo a futuro.

1.8- Organización de la tesis

El desarrollo de esta tesis está dividido en cinco capítulos adicionales al presente: marco teórico, análisis y diseño, implementación, análisis de resultados y, por último, las conclusiones.

El capítulo dos contiene el desarrollo del marco teórico, donde se despliegan los términos e investigaciones que sustentan las bases teóricas de la tesis, entre las que destaca la metodología para la construcción de un modelo ontológico. Además, se muestran algunas técnicas de PLN para análisis de textos en el idioma español en conjunto con algoritmos de clasificación que se utilizarán para la detección de las relaciones entre los elementos de la instrumentación.

En el tercer capítulo, se analiza la arquitectura de la plataforma, así como los pasos de la metodología para la construcción de la ontología en base a los elementos que conforman una ID. A la par, se analiza a profundidad el flujo que seguirán los datos en el módulo de análisis haciendo mención únicamente de la captura de la instrumentación por parte del docente en la plataforma y la retroalimentación del módulo de recomendación. Posteriormente, se analiza el diseño del pseudocódigo para el algoritmo de PLN con el cual se encontrarán las relaciones en base a los patrones del texto redactado.

Para el cuarto capítulo, se detalla la implementación de la ontología utilizando el Software Protege, además de la codificación del algoritmo de PLN realizada en el lenguaje de programación Python. También, se presentan las consultas al modelo de conocimiento empleando el lenguaje ontológico SPARQL.

En el capítulo cinco se analizan los resultados de las pruebas realizadas a la ontología y al algoritmo de PLN utilizando ID redactadas por docentes de educación superior.

Por último, el sexto capítulo contiene las conclusiones del trabajo de tesis en conjunto con las ideas de trabajo a futuro.

Capítulo II-Marco Teórico

2.1- Introducción

En el presente capítulo, se analizan los aspectos teóricos que fundamentan el trabajo de tesis conectando los avances tecnológicos que han permitido a las Tecnologías de la información y comunicación (TIC) adaptarse a los modelos educativos actuales como una herramienta didáctica, ofreciendo beneficios tanto al docente como al alumno que les utilice.

En un principio, se analiza el Modelo Educativo Basado en Competencias (EBC) y se compara con el Modelo Educativo Tradicional (ET) haciendo un cotejo entre ambos y la importancia de un perfil pedagógico para que se aprovechen los recursos didácticos disponibles para el proceso de enseñanza-aprendizaje. Consecuentemente, se muestran las secciones que conforman una instrumentación Didáctica (ID), así como la importancia de una redacción estructurada al momento de elaborarla. También, se documenta la taxonomía de Bloom que fue desarrollada a mediados del siglo XX y es parte fundamental para el funcionamiento del proyecto planteado.

Pasando del tema educativo al tecnológico, se evidencia lo relacionado a ontologías, desde su definición y componentes, hasta los tipos y herramientas más utilizadas para su construcción. A la par, se analizan los modelos de conocimiento basados en reglas, así como las metodologías para su construcción.

Finalmente, se considera lo relacionado a el Procesamiento de Lenguaje Natural (PLN), desde sus niveles de análisis hasta técnicas de recuperación y extracción de información, teniendo en consideración las librerías utilizadas en dichas tareas. En este apartado, se documentan algunos de los algoritmos de clasificación y funciones de similaridad para el análisis de textos.

En resumen, se presentan clasificaciones y terminologías de los principales temas abordados durante la tesis, que dan sustento a la metodología y al análisis de los resultados.

2.2- Modelo Educativo en México

2.2.1- Modelo EBC

El Método ET en el cual el estudiante asistía a clase a memorizar las definiciones, pasos para resolver un problema matemático o un procedimiento para una práctica de laboratorio teniendo al docente como el principal actor frente al grupo, se convirtió en un método obsoleto que no lograba cumplir con las necesidades de la educación actual [4][6][15]. Ante la insuficiencia, a nivel mundial surgió la necesidad de un nuevo paradigma educativo acorde a la demanda laboral regional o local, el cual no compensará la memorización para la acreditación de una materia, si no que permitiera desarrollar ciertas habilidades y aptitudes que garantizaran conocer la utilidad de los conocimientos adquiridos en el aula [6].

Como resultado, el cambio en el paradigma educativo en México se dio en el año 2008 aunque fue hasta el año 2009 cuando se comenzaron a actualizar los planes de estudio adaptándose al Modelo EBC [11]. En él, se buscaba desarrollar habilidades, actitudes, valores y experiencias en el alumno con el fin de poder aplicar la teoría en problemas reales. También, se le motivaba a ser autodidacta, habilidad indispensable en su proceso formativo a lo largo de la vida tanto en el ámbito académico como laboral [6][15].

En este nuevo enfoque, el docente es un asesor que orienta al estudiante y estimula su capacidad autodidacta para el aprendizaje, lo cual conlleva a una nueva modalidad de

formación orientado a un aprendizaje contextualizado para la construcción de conocimiento sin dejar de lado el desarrollo de cualidades y aptitudes por parte del estudiante [4][15].

2.2.2- Puntos del Modelo ET y Modelo EBC

El cambio en el paradigma educativo es un largo proceso que necesita de la formación del docente para facilitar el proceso de enseñanza-aprendizaje, ajustando la forma de generar aprendizaje significativo haciendo uso de las herramientas didácticas disponibles a su alrededor [4][16]. Con esto, se espera que se deje atrás el paradigma del Modelo ET para innovar y poder cumplir con las competencias educativas y digitales [6]. No obstante, algunos maestros tienden a tener una mala práctica al utilizar ambos modelos a la par lo cual, en vez de beneficiar la enseñanza, causa una deficiencia en el proceso educativo [16].

Del mismo modo, la incertidumbre ante el desconocimiento de la metodología correcta para aplicar el Modelo EBC, así como la complejidad y ambigüedad de sus elementos, del concepto de competencia y la falta de conocimiento pedagógicos [7][9], han retrasado la adecuada integración de este modelo educativo, dando resultados muy pobres en comparación con lo esperado a pesar de haber pasado más de una década desde que comenzó su implementación en el país [9][16].

Para entender el cambio del paradigma tradicional al de competencias y algunas de las problemáticas a las cuáles se enfrentan los docentes que se basan en el Modelo ET, es necesario hacer una comparación entre las deficiencias y virtudes que ambos nos presentan. Por eso, en la **Tabla 1** se tiene una comparativa entre los principales puntos en los que está basado el Modelo ET y el Modelo EBC, la cual denota que el Modelo EBC tiene más sentido en la actualidad al ayudar a mejorar el desempeño del estudiante dándole la oportunidad de ser autodidacta, de aprender a aprender y de aprovechar los recursos digitales, ubicándolo como el protagonista del proceso de aprendizaje a diferencia del Modelo ET que opta por tener al profesor como el principal [4][6][16].

Tabla 1-Comparativa del Modelo ET y el Modelo EBC

MET	MEBC
Se tienen fragmentos del mismo tema en diferentes asignaturas, pero estos son vistos sin relacionarse.	Se vinculan las materias entre sí para que el conocimiento adquirido sea de carácter interdisciplinario.
Los temas de las asignaturas son explicados en clase por el docente, mientras que el alumno realiza ejercicios del tema para reforzar conocimiento.	Se busca la resolución de problemas reales utilizando los temas abordados en la materia, trabajando la teoría a la par con la práctica y se muestran distintos enfoques para otorgar un mayor grado de aplicación al tema, de preferencia usando el contexto regional.
La evaluación de la asignatura es en base a exámenes de contenido sin tomar en consideración otros parámetros para la evaluación final.	La evaluación es en base a evidencias del desempeño del estudiante y no en base a la calificación de un examen, lo que permite que el estudiante acredite demostrando que domina el tema y no mediante la memorización.
El proceso de aprendizaje es dirigido por el profesor de principio a fin, siguiendo metodologías de enseñanza que considera las más adecuadas para cumplir con el temario de la asignatura.	Se busca que el alumno sea autodidacta y este en un mejoramiento continuo, dándole un grado de libertad para buscar alternativas de aprendizaje ante las metodologías empleadas por el profesor.

2.2.3- Principio pedagógico del Modelo EBC

Si bien, el Modelo EBC no tiene al docente como el núcleo central, su éxito depende en gran medida de su capacidad para desempeñarse en el aula donde la formación pedagógica juega un papel crucial para guiar el conocimiento [12]. Por ello, es importante que se conozca y se desarrollen los temas de la asignatura de una manera crítica y formativa haciendo énfasis en su comprensión, con el fin de aplicarlo a problemas reales de interés para el estudiante en función del perfil de egreso deseado [5][12].

En este punto, la capacitación y formación son importantes para poder diseñar una ID adecuada con la que se puedan cumplir las competencias planteadas [8][12]. Entre los principales principios pedagógicos que se deben de cumplir para poder consumir las competencias planteadas de la asignatura de acuerdo con el Modelo EBC, se encuentran las siguientes [6][8]:

- El docente debe de ser un facilitador del proceso de aprendizaje, motivando al alumno al desarrollo de habilidades y aptitudes de forma autodidacta. También, debe de preparar el entorno de aprendizaje con estrategias de enseñanza que consideren el contexto.
- El proceso de enseñanza debe centrarse en el desarrollo de competencias enfatizándose en problemas reales y aplicaciones del contexto de acuerdo con el perfil de egreso deseado.
- La evaluación es necesariamente formativa de acuerdo con las competencias establecidas, tomando en cuenta los logros del estudiante y el desarrollo de sus habilidades durante el curso.

2.2.4- Instrumentación Didáctica

La evolución de la educación en México trata de encaminar el aprendizaje secuencial hacia el desarrollo de habilidades y aptitudes que conllevan a un aprendizaje que se evalúa en base al dominio del uso del conocimiento para poder aplicarlo en problemáticas cotidianas, cumpliendo con competencias planteadas como metas [4][6].

Sin embargo, la comprensión del término competencia por sí mismo es complejo[9], debido a que cada docente lo entiende a su manera en base a las experiencias de redacción de años anteriores, acumulando errores de comprensión sin percatarse de su existencia. El desconocimiento de una metodología clara que permita la realización de una ID que relacione las actividades de enseñanza y aprendizaje para posteriormente hacer una evaluación del desempeño del estudiante haciendo uso de un instrumento adecuado, han retrasado significativamente los resultados del Modelo EBC [4][8].

Además, el uso de las TIC en los procesos de enseñanza ha provocado el desarrollo de un nuevo tipo de competencia llamada Competencia digital[3] que busca asociar las tecnologías digitales con el proceso de enseñanza-aprendizaje, el cual es el medio que facilita el conocimiento de los estudiantes y son utilizadas en esencia para motivar su participación, influenciado en gran medida por el perfil de egreso esperado por el alumno [8].

Como base, una ID tiene 4 secciones importantes para definir las actividades que se utilizarán a lo largo del curso, y el cómo estas serán evaluadas. Dichas secciones son las siguientes [8][11]:

- **Actividades de Aprendizaje:** Procedimientos o actividades realizadas por los alumnos para participar en el proceso de formación con el fin de reforzar los conocimientos proporcionados en clase o, de forma autodidacta, adquirir nuevos.
- **Actividades de Enseñanza:** Actividades, técnicas, métodos o procedimientos que el docente utiliza para conducir el proceso de enseñanza-aprendizaje. También se le conoce como estrategias instruccionales.
- **Evidencias de Aprendizaje:** Producto de forma individual y/o grupal que demuestran los resultados del proceso de aprendizaje por parte del alumno.
- **Instrumentos de Evaluación:** Herramientas utilizadas para evaluar las evidencias del desempeño del estudiante entregadas durante el proceso de enseñanza-aprendizaje.

En la secuencia de estrategias que se utilizan para llegar desde las actividades de enseñanza hasta los instrumentos de evaluación, es vital que tengan una congruencia para alcanzar a cumplir las competencias planteadas para la asignatura [9]. Por lo tanto, es importante que el docente comprenda la importancia de estas relaciones y a la vez, que cuente con una profesionalización pedagógica que le permita hacer uso de la metodología para una correcta redacción de su ID acorde al paradigma del Modelo EBC. Como apoyo, se han realizado distintas clasificaciones sobre objetivos de aprendizaje para el diseño de estrategias de enseñanza, siendo una de las más utilizadas la Taxonomía de Bloom[17].

2.2.5- Taxonomía de Bloom

A mediados del siglo XX, el psicólogo y pedagogo Benjamín Bloom propuso un modelo cognitivo de seis partes centrado en las habilidades de pensamiento de una persona. En la taxonomía, se aumentaba de complejidad conforme se iba avanzando de las habilidades de pensamiento de orden inferior a las de orden superior, lo que permitía estructurar el proceso de aprendizaje en base al dominio cognitivo[17].

Los seis niveles que comprenden a la Taxonomía de Bloom para el diseño de estrategias de enseñanza son los siguientes [17][18]:

- **Conocimiento:** Hace referencia al conocimiento de hechos, métodos, teorías, entre otros. Son los elementos que se memorizan.
- **Comprensión:** El alumno es capaz de comprender y aprender la información de tal forma que puede expresarla con sus propias palabras y relacionarlas con otro material.
- **Aplicación:** Habilidad cognitiva que hace referencia al uso práctico o particular del conocimiento.
- **Análisis:** Es considerada una habilidad cognitiva superior que conlleva a descomponer el problema para analizarlo por partes y así, poder relacionar cada una de ellas con el fin de llegar a una conclusión.
- **Síntesis:** Habilidad del pensamiento que se define como la capacidad de un individuo para organizar y relacionar varios fragmentos de conocimiento en un todo que sea coherente.
- **Evaluación:** Habilidad cognitiva que utiliza el juicio para analizar y posteriormente determinar criterios que permitan evaluar el propósito de la actividad a realizar.

En sí, la idea principal de la Taxonomía de Bloom es que, en base a un dominio cognitivo, se tengan relaciones con una lista de verbos que, a su vez se entrelaza con listas de actividades que se pueden utilizar en el proceso de aprendizaje, de tal forma que en conjunto siga una coherencia llevando desde el conocimiento hasta la evaluación [14][17].

2.3- Modelos de Conocimiento

En distintas disciplinas tales como Matemáticas y Física es común encontrar modelos que describen lo que se está trabajando para facilitar el entendimiento del problema, modelando las características para determinar el comportamiento de lo que se está estudiando y poder hacer una predicción ya sea, sobre el mismo modelo o el comportamiento de alguno similar. En consecuencia, un modelo puede considerarse como *“Una representación simple de una realidad compleja, donde se abstraen las características más importantes para comprender la realidad del dominio de estudio”* [19].

Por otro lado, el conocimiento permite comprender sobre lo que se está tratando para poder razonar con el modelo. Partiendo de eso, es posible definir el conocimiento como *“Una acción o efecto de conocer, que se relaciona directamente con el entendimiento y las nociones inscritas a la ciencia y/o sabiduría”*[20].

En conjunto, podemos considerar que los modelos de conocimiento tratan de identificar y describir un dominio en específico, con un panorama general del conocimiento del dominio. Utilizándolos, se pueden construir sistemas informáticos capaces de usar ese conocimiento para resolver problemáticas sin la necesidad de un experto, utilizando reglas para formalizar el conocimiento y encontrar patrones para dar con la solución de la problemática planteada [20][21]. En otras palabras, dichos sistemas informáticos usan el conocimiento y razonan con él utilizando la lógica, reglas de producción, reglas semánticas, lógica descriptiva y otros axiomas que le hacen posible establecer reglas a cumplir para realizar una acción [21].

2.3.1- Sistemas basados en conocimiento

La Ingeniería del Conocimiento es una disciplina encargada de estudiar las maneras de capturar y representar el conocimiento humano [21]. De esa área se desglosan los sistemas basados en conocimiento, los cuáles tratan de imitar lo que haría un humano para resolver una problemática en específico, por lo que comúnmente también son llamados Sistemas expertos [22].

Visto desde una perspectiva general, un sistema experto está compuesto por los siguientes componentes[21]:

- **Base de conocimiento:** Está formada de conocimiento y reglas que le permiten comprender y resolver problemáticas de un dominio en específico.
- **Base de hechos:** Memoria de trabajo que contiene datos específicos sobre los problemas que se desean tratar.
- **Componente de explicación:** Explica el comportamiento del sistema experto para que el usuario pueda comprender el razonamiento que el sistema siguió para dar con la solución del problema.
- **Motor de inferencia:** Interpreta las reglas al proveer metodologías para el razonamiento en base al conocimiento. En él, se interpreta el problema y se mantiene el control utilizando métodos de encadenamiento hacia adelante o hacia atrás para relacionar la entrada con posibles soluciones o viceversa. A la par, se intenta mantener una representación lógica de las soluciones encontradas.
- **Interfaz de usuario:** Facilita la interacción entre el sistema experto y el usuario final. En este apartado el usuario puede hacer preguntas y obtener una respuesta.

El uso de estos modelos puede abarcar múltiples áreas. La unión del conocimiento de varios expertos, así como la simulación del razonamiento humano mediante reglas o modelos probabilísticos, permite tomar decisiones que pueden ser inclusive más eficaces que las de los propios expertos debido a que puede analizar múltiples variables al mismo tiempo e interrelacionarlas para tomar una decisión final [21][22]. En la actualidad existen 3 tipos diferentes de modelos de conocimiento, los cuáles son [22]:

- **Basados en reglas:** Basan su trabajo en la comparación de resultados y aplicación de reglas lógicas para dar una solución a la problemática. En este apartado, entra la regla de *si...entonces...* que es de las más utilizadas en estos sistemas.
- **Basado en casos:** Se busca la resolución de un problema tomando como punto de partido los problemas solucionados anteriormente, es decir, basa su razonamiento en experiencia previa.

- **Basado en redes bayesianas:** Es un modelo probabilístico que representa un conjunto de variables aleatorias y sus dependencias condicionales. En sí, está basado en estadística y en el teorema de Bayes.

2.3.2- Razonamiento

Para modelar un sistema basado en conocimiento, es necesario conocer las reglas sobre las cuáles se basará el razonamiento para resolver la problemática siguiendo una lógica en sus formalismos de representación [22]. Sin embargo, existen dos dificultades que obstaculizan el desarrollo de una buena base de conocimiento: la representación del conocimiento (como codificar el conocimiento para que la computadora pueda entenderlo) y la adquisición de conocimiento [21].

Por lo tanto, tener una base de conocimiento eficiente permite generar reglas más asertivas que representen la heurística del sistema y puedan resolver la problemática al tomar decisiones ideales, simulando la experiencia de un experto en el área mediante el razonamiento realizado por el motor de inferencia [22]. El conocimiento sobre el cual se basa el sistema puede ser [21]:

- **Tácito:** Es subjetivo de cada persona y proviene de buenas prácticas. Entre las técnicas principales se encuentran las reglas de producción, redes semánticas, mapas conceptuales, mapas mentales y diagramas de flujo.
- **Explícito:** Conocimiento documentado que puede ser consultado en caso de que se requiera. Pueden ser teorías, modelos o leyes de acuerdo con la problemática que se desea resolver.

El razonamiento en este tipo de sistemas se da en el motor de inferencia, que es donde se busca la solución a las respuestas del problema. Para llegar a esta inferencia, existen dos estrategias de razonamiento globales [21]:

- **Encadenamiento hacia adelante:** Va desde los datos hacia el objetivo o problema. En otras palabras, a partir de los hechos hace un diagnóstico.

- **Encadenamiento hacia atrás:** Es el caso contrario al partir desde el objetivo o problema hacia los datos, por lo que va desde la hipótesis en busca de los hechos que han derivado el problema.

En la práctica, es habitual encontrar un híbrido entre ambos encadenamientos ya que no siempre se puede hacer un diagnóstico a partir de los hechos o viceversa, por lo que en conjunto dan como resultado un sistema más completo [21].

2.3.3- Ontologías

El término ontología es un concepto que surgió en la antigua Grecia derivado de las palabras *ontos* del *ente* o *ser* y *logos* que es *teoría* o *palabra*[23]. En aquella época estaba estrechamente relacionada con la Filosofía, pero a principios de la década de los noventa se fue ligando con otras áreas de manejo de información tales como la Inteligencia Artificial (IA) y el Procesamiento de Lenguaje Natural (PLN) [19].

El uso de este término se incluyó dentro del área de las Ciencias Computacionales ante la necesidad de tener una correspondencia entre las entidades de un dominio tomando en cuenta sus propiedades para representar conocimiento en conjunto con las relaciones entre los elementos que integran el área. El objetivo central era facilitar la recuperación de información gracias a la liga existente entre los elementos [19].

Por este motivo, en el año de 1998 Studer [23] definió una ontología como “*Una especificación explícita de una conceptualización*”, siendo esta una de las definiciones más aceptada dentro del área computacional de su época. Desde entonces, se han publicado definiciones más actuales para este término, pero todas se derivan de la representación de conocimiento y la relación entre los conceptos.

Generalmente, las investigaciones realizadas alrededor de ellas están relacionadas con la web semántica o también llamada web 3.0 [23]. La principal razón de este nombramiento radica en que los modelos ontológicos permiten almacenar información de uno o diversos dominios e interrelacionarlas [19].

La tendencia de la web está encaminada hacia el uso de algoritmos para transformar la información en conocimiento con el objetivo de que la computadora pueda entender los datos sin que un humano esté supervisando. La idea central está basada en el uso de metadatos que contengan información adicional a la que se muestra en formato texto, donde se expresen características que describan información adicional para inferir conocimiento [19]. Ante esta necesidad, las ontologías fueron anexadas a la informática con la intención de representar conocimiento y deducir nuevo a través de ellas, tratando de llevar la información almacenada a una representación que permitiera su recuperación y extracción de una manera eficaz[23].

2.3.4- Componentes de una Ontología

Para la representación del conocimiento, las ontologías contienen algunos componentes claves que le permiten estructurar los datos e interrelacionarlos entre sí para inferir conocimiento. A grandes rasgos, su base es la relación entre las entidades (dominio a modelar) y los atributos (descripción de las entidades), donde se tienen axiomas (condiciones lógicas) a cumplir. Sus principales componentes son los siguientes [11][19]:

- **Conceptos:** Son el dominio del modelo a modelar, las ideas básicas que se intentan formalizar. También son llamados clases y tienen asociados atributos que los describen. Un ejemplo de ello sería una ontología escolar sobre *materias*, donde *ciencias* sería una clase.
- **Atributos:** Describen más detalladamente las clases y sus instancias. Por ejemplo, una *materia* puede ser *aprobada* o *reprobada*. En cuanto a los atributos de la instancia, estos pertenecen a cada instancia y se pueden heredar a las instancias y subclases. Un ejemplo de ello es el *número identificador* de la materia o el *número de control* de un alumno.
- **Instancias:** Son representación de individuos u objetos precisos de un concepto en una ontología. Por ejemplo, *Cálculo* y *Álgebra* son instancias de la subclase *Exactas*, ya que pueden interpretarse como materias derivadas de las ciencias exactas.
- **Relaciones:** Son los enlaces y relaciones que se tienen entre dos clases del dominio o entre los conceptos del dominio y suelen tener propiedades que describen su

comportamiento para construir taxonomías del dominio que se está modelando. Ejemplos de relaciones pueden ser *subclase-de*, *parte-de* y *pertenece-a*.

- **Funciones:** Es una relación definida que identifica un elemento mediante el cálculo de una función que considera varios elementos del dominio. Por ejemplo, un modelo que contenga las *notas y materias* de un *alumno*, se pudiera tener funciones del tipo *sumar-materia*, *promediar-materia*, entre otras funciones.
- **Axiomas:** Son teoremas y expresiones lógicas que deben de cumplirse entre los elementos para definir restricciones en ella en base a valores específicos de sus propiedades. Por ejemplo, no se puede cursar *Cálculo I* y *Cálculo II* durante el mismo semestre. A su vez, los axiomas también suelen ser llamados reglas de restricción, y estas sirven para deducir conocimiento entre los atributos e instancias.

2.3.5- Tipos de Ontologías

Su clasificación es transformada dependiente del autor y varía en cuanto a los niveles y el tipo de conceptualización abordada en ella. Por ejemplo, en el año de 1997, Van Heijst [19] propuso una clasificación basada en la estructura de la conceptualización y la base del tema de la conceptualización, donde las dividía en terminológicas, información y modelado de conocimiento. De las distintas clasificaciones que se tienen, la mayoría coinciden en que las Ontologías están divididas de la siguiente manera [11][19] :

- **De dominios:** Representa un modelo de conocimiento especializado de un dominio o subdominio en particular. Cumple con la generalización de tareas concretas en un dominio específico de conocimiento. Por ejemplo, puede ser un modelo de conocimiento sobre *mecánica*, *eléctrica*, *electrónica*, entre otras.
- **Genéricas:** Son conceptualizaciones muy básicas y generales entre los dominios y son consideradas independientes.
- **De aplicación:** Están ligadas a un dominio y tarea en particular, lo cual genera una especialización de esta.

2.3.6- Metodologías para Ontologías

Los pasos para diseñar cualquier ontología deben de ser muy específicos, por lo que existen diversas metodologías que se pueden utilizar de acuerdo con el tipo que se pretenda construir [24]. Así pues, el paso principal que está disponible en todas las metodologías existentes es la extracción del conocimiento desde diversas fuentes para, a partir de él, inferir nuevo conocimiento. Desde esa iniciativa, se desarrollaron algunas metodologías para su diseño entre las que destacan las siguientes [19][24]:

- **USCHOLD y KING:** Desarrollada en 1995, contiene una serie de pasos que permiten identificar el propósito, capturar los conceptos y relaciones con los términos utilizados para referirse a ellos, para finalmente codificarlos. The Enterprise Ontology (W3C) es un proyecto desarrollado con esta metodología y su información es sobre empresas de negocios.
- **KACTUS:** Fue desarrollada en el año de 1996. La construcción es en base a un dominio y, a partir de él se van estructurando y adaptando a nuevos.
- **METHONTOLOGY:** Es de las metodologías más completas que existe para su construcción. Fue desarrollada en 1996 y se compone de 5 etapas: especificación, conceptualización, formalización, implementación y mantenimiento.
- **ON-TO-KNOWLEDGE:** Hace uso de la información digital disponible, siendo esta una de sus principales ventajas respecto a las otras ya que, a partir de esa información recabada gestiona el conocimiento basada en el análisis de escenarios de uso.
- **TERMINAE:** Desarrollada en el año 2002, permite la construcción a partir de textos. Se dividía en dos herramientas principales: *Syntex* que permitía identificar los términos y relaciones; mientras que *Caméléon*, es utilizada para identificar los roles y las relaciones.

2.3.7- Lenguajes y herramientas para Ontologías

Existen varios lenguajes ontológicos que fueron desarrollados para codificarlas de tal forma que la computadora pudiera entenderlos. No obstante, cada uno tiene sus ventajas y desventajas respecto a otros, por eso se debe de tener en cuenta la aplicación a desarrollar

para elegir entre cada uno de ellos. Entre los principales lenguajes ontológicos desarrollados se encuentran los siguientes [19]:

- **CYCL**: Es un lenguaje cuya sintaxis se deriva del cálculo de predicciones de la lógica primer grado y está basado en marcos.
- **ONTOLINGUA**: Se utiliza para expresar conocimiento y está basado en aproximaciones de marcos y lógica de primer orden, utiliza términos tales como clases, subclases, instancias, relaciones, funciones, objetos y axiomas que relacionan términos entre sí, con lo cual pudo resolver problemas de otros lenguajes en cuanto a la representación del conocimiento.
- **RDF**: Este modelo equivalente a una red semántica de representación de conocimiento para describir recursos de la web. Su sintaxis es en formato XML en la cual sus nodos representan conceptos, instancias de conceptos y el valor de sus propiedades. Junto a OWL es uno de los más utilizados.
- **OWL**: Este lenguaje surge en el 2004 con la búsqueda de un lenguaje estandarizado para los investigadores de la web semántica. En él, es posible definir clases mediante restricciones u operaciones sobre otras clases. Así mismo, se tiene una relación entre las clases como la inclusión, disyunción y equivalencia, con lo que se pueden definir restricciones de cardinalidad de las propiedades al componerse de una secuencia de axiomas y referencias de otras ontologías al ser documentos web referenciados a través de una URI.
- **Protégé** [25]: Herramienta cuya interfaz es amigable e intuitiva. Permite la construcción de modelos de conocimiento de dominio y exportar el archivo resultante en formato OWL o RDF. En sí, el modelado es mediante la creación de clases, instancias, funciones, relaciones, entre otras. Uno más de sus puntos a favor es que cuenta con módulos para la visualización de las relaciones y para la realización de consultas utilizando el lenguaje SPARQL, complementando su entorno con un amplio catálogo de plugin con diferentes funcionalidades.

2.4- Procesamiento de Lenguaje Natural

El Procesamiento de Lenguaje Natural (PLN), es una disciplina que basa sus estudios en la comunicación entre una computadora y un ser humano teniendo como intermediario el lenguaje natural, ya sea de forma oral o escrita [26]. Una definición teórica actualizada aplicable a distintas áreas enfocadas en el uso del PLN es: *Aplicación multidisciplinar de técnicas de tratamiento de lenguaje para la resolución de problemas de diversas áreas, que necesitan de la interpretación del lenguaje humano, por medio de técnicas y métodos computacionales* [27].

Al ser de carácter multidisciplinar, el PLN tiene múltiples aplicaciones al poder abarcar cualquier área en la cual se tenga texto de por medio [27]. En un inicio, las principales aplicaciones a las que estaba destinado el PLN era a la traducción de textos, la recuperación y extracción de información y, por último, al desarrollo de interfaces humano-máquina [28].

Por otra parte, con los avances en las investigaciones del área, las aplicaciones del PLN se han extendido a otros temas. En particular, con la integración de la IA, los algoritmos de PLN han permitido tener sistemas de búsqueda de respuestas más aptos y precisos, analizar los sentimientos de un texto, la generación automática de resúmenes y la minería de textos. Por estos avances, las áreas de aplicación van desde la medicina, hasta el uso en la educación o atención a personas para satisfacer la necesidad de información [29]. Como consecuencia, el PLN se utiliza diariamente sin tenerlo en cuenta, ya sea cuando se utiliza el buscador en la web, o con las predicciones que hace el teclado al momento de escribir.

Partiendo de este amplio catálogo de aplicaciones disponibles en la actualidad anexado a la digitalización de la información existente en papel, se ha dado un incremento exponencial de los datos almacenados en la web, generando volúmenes de información que está en su mayor parte dispersa en la nube sin metadatos que permitan su recuperación con técnicas convencionales [30]. Para entender esta problemática, es importante destacar que la Recuperación de Información es un proceso que busca en grandes colecciones de datos de tipo estructurada (base de datos), semi estructurada (como la web) o en forma no estructurada (texto en formato libre), la información necesaria que satisfaga la necesidad de esta [29].

Cabe resaltar que, se debe de hacer una distinción entre la recuperación y la extracción de información. La extracción, es la obtención de las partes de un texto y se utiliza con la finalidad de pasar esa información a un formato estructurado, mientras que la recuperación es la búsqueda de esa información en las colecciones de datos. Para medir la efectividad de ambas tareas, se utilizan dos medidas [29]:

$$\text{cobertura} = \frac{\text{numero de documentos recuperados}}{\text{numero total de documentos relevantes}}$$

$$\text{precision} = \frac{\text{numero de documentos relevantes recuperados}}{\text{numero total de documentos recuperados}}$$

En repositorios muy grandes, a veces es primordial tener una alta precisión aun cuando la cobertura no sea buena, mientras que, en otras se sacrifica precisión a cambio de más documentos recuperados. La elección de uno u otro dependerá en gran medida de la aplicación a la cual este destinado el sistema y que sea más importante, si tener una mayor cantidad de documentos recuperados aun cuando estos tengan poca relevancia ante la consulta o, el tener solo unos cuantos documentos, pero el grado de relevancia sea lo suficientemente alto para asegurarse de satisfacer la necesidad de información.

Como se ha expresado, la tendencia al alza en cuanto a la información digital ha provocado que empresas tales como IBM realicen estimaciones de la cantidad de información digital disponible en la web. En el año 2012, IBM [30] indicaba que, alrededor del 90 % de los datos digitales se habían digitalizado en los últimos dos años, dando una aproximación de 2.5 trillones de bytes de datos (1 exabyte) de datos creados diariamente, siendo las redes sociales las plataformas que más aportan a esta estadística.

Otra estimación realizada en noviembre del año 2018 en un informe [31] publicado por la empresa IDC en conjunto con SEAGATE menciona que, para el año 2025 se tendrán 5 veces más datos de los existentes en el año 2018, pasando de 33 Zettabits de datos almacenados a 175 Zettabits. Por estos motivos, el amplio volumen de datos creados diariamente denota que es una tarea imposible su etiquetado manual y el almacenarlos de manera estructurada para facilitar su recuperación [30].

En última instancia, sobre este tipo de predicciones recae la importancia de las tareas de PLN para el análisis de textos en formato libre, pudiendo hacer recuperación de la información en grandes volúmenes de datos aun cuando estos se encuentren en forma no estructurado [30], por lo que la integración de algoritmos basados en aprendizaje maquina y los avances en ciencia de los datos han permitido mejoras en su precisión, además de ayudar en la identificación de patrones en el texto que a simple vista pueden parecer inexistentes.

2.4.1- Niveles de análisis y Normalización

Para el análisis computacional de una frase utilizando PLN, se requiere tener en cuenta las bases del conocimiento de una lengua (fonética, morfología, sintaxis, semántica y gramática de contexto). Para sistematizar este conocimiento, es necesario hacer uso de una serie de pasos o niveles de análisis del PLN, los cuáles son [27]:

- **Morfológico:** O nivel léxico. En él se realiza una segmentación de la frase en tokens (palabras) que son lematizados (raíz del token) y analizadas como elementos únicos con el fin de extraer la categoría sintáctica y su significado léxico.
- **Sintáctico:** También denominado Parsing. Convierte las frases en unidades léxicas de tipo árbol, donde se encuentra el papel estructural de cada palabra dentro de la frase, así como su relación jerárquica.
- **Semántico:** Este nivel parte del nivel sintáctico y trata de encontrar el significado de las palabras y el sentido de la frase. Es uno de los niveles más complicados de trabajar debido al problema de la de ambigüedad, además de que algunas palabras tienen significado diferente dependiendo la región del mundo donde se estén utilizando.
- **Pragmático:** En este nivel, la computadora trata de comprender e interpretar la frase para darle un significado asertivo en base al contexto.

Los niveles de análisis van relacionados entre sí para poder darle sentido a una frase computacionalmente. Partiendo de la segmentación de frases, se trata de buscar el papel estructural (si la palabra es verbo, sustantivo, adjetivo, entre otros) dentro de la oración, para posteriormente encontrar su significado en base al contexto de uso con el fin de otorgar una interpretación final. Por otra parte, no siempre se usan todos los niveles de análisis para la

solución de un problema de lenguaje natural y cuáles se utilizarán depende que tan profundo sea el análisis que se desea realizar.

En conjunto con los niveles de análisis, se realizan otro tipo de tareas para normalizar los textos y así, facilitar el proceso de análisis. Entre las principales tareas de este tipo, podemos encontrar las siguientes [26]:

- **Tokenización:** Segmentación o división de un texto en tokens (cadena de caracteres con significado dentro del contexto del texto).
- **Segmentación:** División del texto en frases o párrafos a fin de poder ser analizados de manera independiente.
- **Etiquetado gramatical:** también llamado POS (part-of-speech) o Tagging, es una asignación de un etiquetado de tipo gramatical (pronombre, verbo, adjetivo, sustantivo, etc.), que indica el rol gramatical de un token dentro de una frase.
- **Reconocimiento de entidades nombradas:** Entidades como nombres de personas, empresas, siglas, fechas o lugares, que dan información por sí mismas.

2.4.2-Similitud entre cadenas

Para el análisis de textos, es importante tener una métrica que permita determinar si dos cadenas son iguales o que tan similares son entre sí. Con ese fin, se han desarrollado funciones de similitud que retornan un valor numérico (entre 0 y 1) que indiquen un mayor grado de similitud cuando el resultado sea cercano a 1. Como complemento, es utilizado el termino distancia como el inverso de similitud ($s = 1-d$), con lo que se concluye que a mayor distancia se tiene una menor similitud y viceversa [32].

Desde un punto de vista genérico, para medir la similitud entre textos existen dos grupos principales: las *medidas compuestas* (medidas basadas caracteres) que segmentan el texto en tokens y calculan la similitud de cada uno para después unificar los resultados obteniendo un único valor; y las *medidas no compuestas* que representan el texto en un modelo abstracto (normalmente vectores) para realizar el análisis utilizando esa representación [33].

A la vez, las medidas compuestas están subdivididas en: *semánticas* que basan su resultado en la comparativa con grandes volúmenes de información para determinar si dos palabras son similares semánticamente utilizando redes semánticas o matrices; y *métricas léxicas* tales como la subsecuencia común más larga, algoritmo que calcula cuán similares son dos cadenas en base a la longitud de la cadena próxima más larga existen entre ambas. Otra medida léxica por considerar es la de Jaro-Winkler, quien compara dos palabras para identificar cuántos caracteres están presentes en ambas, considerando tanto la posición como si están transpuestos [27][33].

Por su parte, una función de similitud común es la distancia de edición o también llamada Distancia de Levenshtein, la cual realiza la comparativa entre dos cadenas (A y B) para determinar cuántas operaciones de edición (insertar, substituir y eliminar caracteres) se necesitan para convertir la cadena A en B o B en A, otorgando un costo por cada modificación que se realiza [32].

Como alternativa, se tienen funciones que utilizan subcadenas segmentadas por caracteres especiales (espacio en blanco, signos de puntuación, entre otros) y en base a ellas se obtiene la similitud entre el conjunto de caracteres. Una de estas métricas es la similitud coseno TF-IDF (Frecuencia de Término-Frecuencia inversa de Documento), la cual ayuda a evaluar la relevancia de una palabra en un texto convirtiendo las cadenas en representaciones vectoriales de n componentes para posteriormente hacer uso de la función coseno con el fin de medir el ángulo que forman los vectores y así obtener el grado de similitud entre ellos [32].

2.4.3- Librerías de PLN

Para el análisis de textos, se han creado librerías tanto de paga como de código abierto con el fin de que la máquina pueda entender el lenguaje natural. En la actualidad, las principales librerías para el PLN han sido desarrolladas para el lenguaje de programación Python¹, uno de los más utilizado en el área de la IA. Entre las librerías más usadas para el análisis de textos en el idioma español se encuentran las siguientes:

¹ <https://www.python.org/>

- **NLTK** [34]: Natural Language Toolkit es una librería de PLN creada en el año 2001 con el fin de analizar textos en inglés. NLTK es de las más utilizadas en el mundo, cuenta con más de 50 corpus y recursos léxicos, así como una amplia variedad de herramientas que permiten tokenizar, etiquetar, hacer análisis semántico y sintáctico. Como punto negativo tiene que solo algunas funcionalidades están disponibles para el español.
- **Freeling** [35]: Librería de código abierto desarrollada en la Universidad Politécnica de Cataluña (UPC). Entre sus principales funcionalidades está el permitir tokenizar, realizar análisis morfológico, detectar entidades nombradas y el etiquetado POS. Todas las funcionalidades son soportadas en el idioma español y es considerada una librería multilinguaje. Su implementación está en el lenguaje C++, pero puede ser utilizada mediante APIs desde otros lenguajes, entre ellos Python.
- **Stanford CoreNLP** [36]: Conjunto de herramientas para PLN diseñada y desarrollada para el lenguaje Java que permiten el etiquetado POS, tokenizado y el reconocimiento de discursos. Algo a considerar es que NLTK hace uso de algunas de las funcionalidades de esta librería y al igual que ella, no todas las funcionalidades están disponibles para el español.
- **Spacy** [37]: Librería de código abierto multilinguaje al soportar más de 45 idiomas, entre los que se incluye el español. En ella, se pueden encontrar modelos estadísticos de palabras que permiten reconocer entidades nombradas, así como un etiquetado POS completo de la funcionalidad gramatical de cada token dentro de una oración. Es una librería completa y se caracteriza por su facilidad de uso.
- **Word2Vec** [38]: Desarrollado en Google por Tomas Mikolov en el año 2013, es un modelo de espacio vectorial de n dimensiones basado en redes neuronales. Word2vec está disponible de dos formas: El Modelo Continuo de Bolsa de Palabras (CBOW) que, a partir de una fuente de palabras, trata de predecir la palabra objetivo. En contraparte, Skip-Gram funciona a la inversa, ya que, a partir de una palabra objetivo, trata de predecir el contexto de las palabras vecinas.
- **BERT** [39]: Bidirectional Encoder Representations from Transformer, es una librería de código abierto desarrollada por Google y es una mejora de Word2Vec que considera al contexto de la palabra para formar su vector, caso contrario a Word2Vec

que solo tomaba la palabra sin considerarlo. Además de ser multilinguaje, su objetivo principal es mejorar la comprensión del lenguaje y así poder entender mejor el contexto de las palabras dada una consulta.

2.4.4- Algoritmos de clasificación

El aprendizaje automático (Machine Learning), es una rama de la IA que permite el reconocimiento de patrones a partir de los datos y es aplicado en imágenes y texto, lo que favorece que pueda extenderse su aplicación a múltiples áreas. Visto de una manera genérica, los algoritmos de aprendizaje automático pueden agruparse de la siguiente manera [40]:

- **Aprendizaje supervisado:** Basándose en ejemplos previamente etiquetados, se busca encontrar la función capaz de mapear la entrada y la salida.
- **Aprendizaje no supervisado:** En este tipo de aprendizaje se busca encontrar patrones en los datos para poder agruparlos y, a partir de eso, etiquetar nuevos datos. El método más utilizado es el clustering.
- **Aprendizaje semi-supervisado:** Combina el aprendizaje supervisado y no supervisado al hacer uso de ejemplos previamente etiquetados y datos sin etiquetar para realizar la clasificación.
- **Aprendizaje por refuerzo:** El algoritmo es capaz de aprender en base a procesos de prueba y error, reforzándose de aquellas pruebas positivas. Es importante tener claras los atributos que se utilizaran para una clasificación, ya que de la elección de dichos atributos dependerá la precisión de esta.

Para elegir, se tienen múltiples algoritmos de clasificación con sus ventajas y desventajas dependiendo de la tarea que se desea resolver. Los algoritmos más comúnmente utilizados son los siguientes[29][40]:

- **K Nearest Neighbor (KNN):** también conocido como vecino más cercano. Es un algoritmo que trata asignar los k vecinos más cercanos para formar una clase. Entre las desventajas de este algoritmo, es que tarda más tiempo clasificando los objetos respecto a otros algoritmos de clasificación.

- **Algoritmo Iterativo de Naïve Bayes:** Algoritmo estadístico que está basado en el teorema de Bayes, cuyo objetivo es encontrar patrones que adecuen el modelo más probable tomando como base documentos previamente etiquetados o sin etiquetar.
- **Support Vector Machines (SVM):** El algoritmo de máquina de soporte vectorial trata de encontrar separación entre las clases de los datos previamente etiquetados, esto en base a un mapeo de los datos de entrada hacia un modelo multidimensional que, mediante el uso de algunos métodos, permita construir un hiperplano para la clasificación de los datos.
- **Clustering:** O también llamado algoritmo de agrupamiento. Su tipo de aprendizaje es no supervisado y en él se tratan de encontrar características y propiedades similares entre los objetos que se pretenden agrupar. Como complemento, la teoría marca que aquellos objetos que no tienen características similares deben de pertenecer a distintos agrupamientos. Este tipo de clasificación está basada en un análisis estadístico de los datos y como tal tiene múltiples aplicaciones en diferentes áreas.

Capítulo III- Análisis y diseño

3.1- Introducción

El Modelo Educativo basado en Competencias (EBC) contiene un extracto con estrategias didácticas para que el docente pueda utilizar como herramienta de apoyo en sus clases. Las tareas que sean utilizadas para cumplir con los objetivos de aprendizaje deben de considerar el perfil de egreso del estudiante enfocándose en que aprenda más allá de la teoría, desarrolle habilidades, aptitudes y valores que estén encaminados a metas educativas que se alcanzan al cumplir la competencia de la materia.

Sin embargo, planear un curso implica que se relacionen los temas con las actividades, evidencias e instrumento quedando todo plasmado en la Instrumentación Didáctica (ID) de la asignatura. Dado que es texto en formato libre, es factible utilizar técnicas de PLN para la detección de patrones y, al ser un dominio muy específico, ontologías para almacenar la información e inferir nuevo conocimiento a partir de las relaciones existentes, siendo esta una de las principales ventajas de su utilización.

Por lo tanto, en el presente capítulo se muestran los diagramas de la arquitectura del sistema enfocándose en el desarrollo del módulo de análisis que contiene el diseño y construcción de la ontología, basándose en la metodología Methontology, recomendada por la Fundación de los Agentes Físicos Inteligentes. Posteriormente, se realiza un análisis del diseño del algoritmo de PLN mostrando diagramas que ayudan en su interpretación, así como una vista del flujo de los datos y la interacción del algoritmo con el modelo de conocimiento.

3.2- Arquitectura del sistema

El proyecto planteado en el presente trabajo corresponde a un algoritmo encargado de evaluar la ID de asignaturas de nivel superior tecnológica utilizando técnicas de PLN y un modelo de conocimiento. El algoritmo es identificado como el módulo de análisis y pertenece a un sistema dedicado a la capacitación y formación docente, cuya arquitectura propuesta se muestra en la **Figura 1** donde se pueden identificar las 3 secciones que lo componen:

- **Plataforma:** En esta sección, el profesor escribirá su ID y podrá consultar el material disponible, además de las sugerencias que sean realizadas por el módulo de recomendación. Cabe resaltar que el diseño, construcción y funcionamiento de la plataforma son ajenos al análisis del proyecto planteado, por lo que su función radica en ser la fuente de datos hacia el módulo de análisis.
- **Análisis:** Es el encargado de examinar los datos de la ID y es sobre el que se basa la presente investigación. En la **Figura 1** se muestra que el módulo está dividido en dos secciones: la primera de ellas se basa en la organización de los datos en la ontología; mientras la segunda parte está relacionado al algoritmo de PLN empleado para el análisis del texto en formato libre. Como resultado del análisis, se obtienen metadatos que son enviados al módulo de recomendación.
- **Recomendación:** Por último, este módulo recibe los datos del módulo de análisis para otorgar una retroalimentación. En dado caso de detectar deficiencias en la ID, se encarga de hacer las recomendaciones pertinentes.

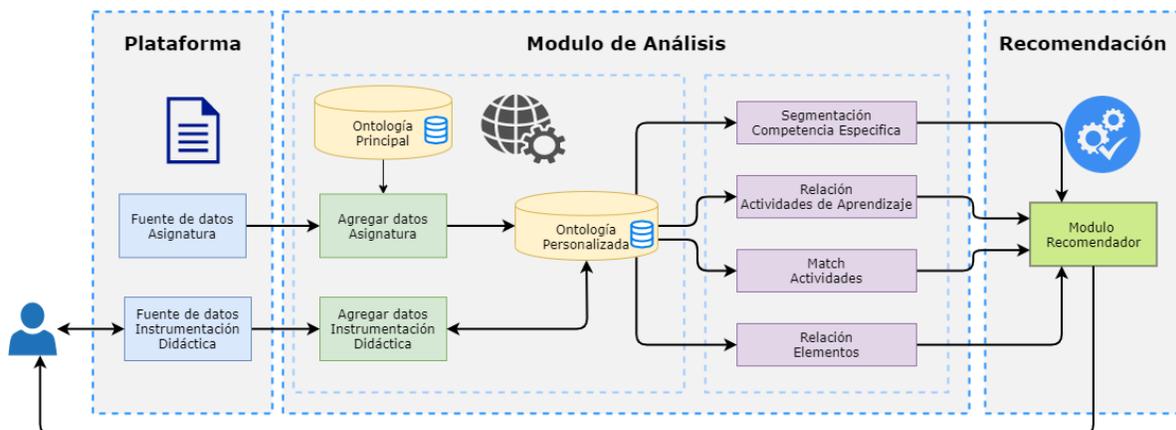


Figura 1-Arquitectura completa del sistema

En conjunto, el sistema se encarga de ayudar al docente en el entendimiento de la metodología para una correcta redacción de su ID, otorgándole material didáctico y una retroalimentación instantánea. Como se mencionó anteriormente, la presente investigación se centra solo en el desarrollo del módulo de análisis, por lo que el diseño, construcción y funcionamiento tanto de la plataforma como del módulo de recomendación no son descritos en las siguientes secciones del presente capítulo.

3.3- Diseño de la base de datos

La plataforma de capacitación contiene algunas tablas que le permiten indicar al módulo de análisis cuál es la ontología personal. Para almacenarla, se tomaron algunas consideraciones en las relaciones como lo son que un docente puede pertenecer a más de una carrera y, por lo tanto, impartir más de una asignatura. En consecuencia, puede tener más de una ID, por lo que se optó por utilizar el RFC y la clave de la asignatura como clave primaria.

También, se cuenta con algunos catálogos que contienen lisados predefinidos que están almacenados para facilitar a la plataforma el acceso a ellos. Ahora bien, para el módulo de análisis, estos catálogos son consultados directamente en la ontología. En la **Figura 2** se muestra el diagrama que contiene las tablas y relaciones existentes en la base de datos de la plataforma, así como la cardinalidad. En cuanto al modelo Entidad-Relación en el cual está basado el diagrama es el siguiente:

- ✓ **Asignatura** = {claveA, nombreA, temario}
- ✓ **Profesor** = {RFC, nombre}
- ✓ **Carrera** = {claveC, nombreC, departamento}
- ✓ **profesorCarrera** = {RFC, claveC}
- ✓ **InstrumentaciónDidáctica** = {claveA, RFC, año, semestre, grupo, modelo, documento, unidades}
- ✓ **catActividadesAprendizaje** = {idActividadA, nombre}
- ✓ **catActividadesEnseñanza** = {idActividadE, nombre}
- ✓ **catEvidenciaAprendizaje** = {idEvidenciaA, nombre}
- ✓ **catInstrumentosEvaluacion** = {idInstrumentoE, nombre}
- ✓ **catCompetenciasGenericas** = {idCompetencias, nombre, tipo}
- ✓ **catIndicadorAlcance** = {idIndicadorA, nombre, valor}

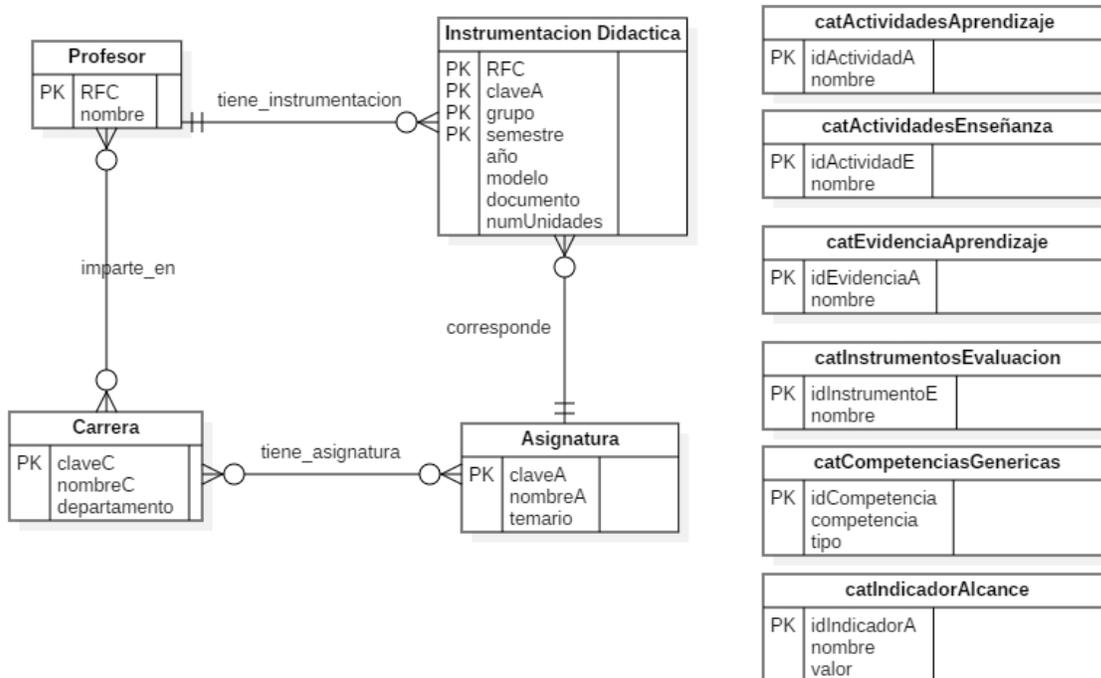


Figura 2-Diagrama de la base de datos del sistema

3.4- Diseño y procesos de la ontología

3.4.1- Metodología Methontology

El diseño y construcción de una ontología requiere del seguimiento de una serie de actividades que permitan modelar el dominio a nivel conceptual a partir de la adquisición de conocimiento, para posteriormente ser transformado a un modelo formal en algún lenguaje ontológico como puede ser OWL o RDF. Una de las metodologías que permite su diseño y construcción partiendo desde un nivel de conocimiento es la metodología Methontology compuesta por 5 pasos esenciales [21]:

- **Especificación:** Limita el dominio a un área específica de conocimiento.
- **Conceptualización:** Mediante la adquisición de conocimiento, permite la creación de glosarios que contienen los conceptos, atributos, taxonomías y sus relaciones.
- **Formalización:** Admite convertir el modelo de conceptualización en un modelo formal utilizando lenguajes ontológicos.

- **Implementación:** Haciendo uso de herramientas tales como Protégé, se convierte el modelo formal en computable.
- **Mantenimiento:** Permite realizar actualizaciones.

Los pasos que seguir de la metodología Methontology son muy específicos, aunque la organización de la secuencia de sus actividades permite partir del conocimiento de un experto o hacer reingeniería para que sea entendida mediante representaciones graficas aun para quienes no son expertos en el área.

El dominio de la presente ontología contiene información referente a la terminología y relaciones encontradas en una ID de asignaturas de nivel superior tecnológica, por lo que su dominio es específico de esta temática. En las siguientes secciones del presente capítulo se describen las tareas de la metodología Methontology a la par con las tablas y diagramas que contienen el conocimiento adquirido para su construcción.

3.4.2- Conceptualización

Una vez que se ha definido el dominio, el siguiente paso es la conceptualización del conocimiento adquirido de expertos en el área para posteriormente ser computarizado utilizando alguno de los lenguajes ontológicos. En la estructuración y modelado del conocimiento recabado, se utilizan tablas y diagramas para almacenar los conceptos, instancias, atributos, axiomas y relaciones entre los términos a modelar, siguiendo una serie de tareas marcadas en la metodología Methontology.

Tarea 1-Glosario de términos: La primera tarea que marca la metodología es la construcción de un glosario de términos que serán utilizados, incluyendo a todos aquellos que son relevantes para el dominio a modelar, así como una descripción de los conceptos tal y como se muestra en la **Tabla 2** (Ver **Anexo 1**).

Tabla 2-Glosario de conceptos y su descripción (Anexo 1)

Conceptos	Descripción
Actividades de aprendizaje	Actividades realizadas por los estudiantes para reforzar los conocimientos proporcionados por el docente.
Actividades de aprendizaje predefinidas	
Actividades de aprendizaje seleccionadas	
Análisis de competencias específicas	Análisis de cada una de las competencias específicas a partir de los cuatro objetos (temas y subtemas, actividades de aprendizaje, actividades de enseñanza, desarrollo de competencias genéricas).

Mientras tanto, los individuos de la **Tabla 3** (ver **Anexo 2**), contiene el glosario de individuos correspondientes a los conceptos. En total, la ontología contiene 83 clases (21 son principales y 69 subclases) y 486 individuos pertenecientes en su mayor parte a los predefinidos para la Taxonomía de Bloom.

Tabla 3-Glosario de individuos y sus sinónimos (Anexo 2)

Individuo	Sinónimos	
Abstraer	Abstracción	
Acordar	-	
Actividad focal inductora	-	
Agrupar	Agrupación	
Aislar	-	
Analizar	Análisis	
Analogía	-	
Aplicar	Aplicación	
Apoyar	-	
Apreciar	Apreciación	
Aprendizaje basado en análisis	Análisis descriptivo Análisis explicativo	Análisis evaluativo Análisis dato

Tarea 2-Jerarquía de clases: Los conceptos de dominio también suelen denominarse superclases. Debido a que algunos son muy generales, se utilizan subclases para tener un concepto más específico dando como resultado una jerarquía (taxonomía). Las clases más importantes de la presente ontología se basan en la Taxonomía de Bloom, por lo que esta tarea permite ordenarlas adecuadamente. Como se mencionó en el paso anterior, se tienen 83 clases, de las cuales 21 son superclases y estas contienen a 62 subclases.

Una vista parcial de la jerarquización es la siguiente:

- Actividades de Aprendizaje
 - Actividades de Aprendizaje predefinidas
 - Actividades de Aprendizaje predefinidas Análisis
 - Actividades de Aprendizaje predefinidas Aplicación
 -
 - Actividades de Aprendizaje seleccionadas
- Actividades de Enseñanza
 - Actividades de Enseñanza predefinidas
 - Actividades de Enseñanza predefinidas Análisis
 - Actividades de Enseñanza predefinidas Aplicación
 - ...
 - Actividades de Enseñanza seleccionadas
- Análisis de competencias específicas
- Apoyos didácticos
- Asignatura
- Competencias
- ...

El resto de la jerarquía de clases está disponible en el **Anexo 3**.

Tarea 3-Atributos de las clases: Los conceptos pueden tener atributos que fungan como variables dentro de la ontología y también suelen ser llamadas propiedades de conceptos. La **Tabla 4** (Ver **Anexo 4**) contiene la relación entre el dominio y las propiedades, así como el tipo de dato que será almacenado.

Tabla 4-Propiedades de las clases (**Anexo 4**)

Dominio	Relación	Propiedad	Tipo
Análisis de competencias específicas	tiene	Descripción	String
		Diagnóstica	
		Fecha de final	
		Fecha de inicio	
		Formativa	
		Nombre	
		Número	
		Sumativa	

Tarea 4-Relaciones entre conceptos: Para poder inferir conocimiento, es necesario tener relacionados los conceptos entre sí estableciendo su dominio y un rango. Una de las ventajas de esta tarea, es que facilita llevar las relaciones hacia una representación visual tal como un diagrama, donde es más cómodo observar las relaciones existentes entre los conceptos. La **Tabla 5** (ver **Anexo 5**) contiene una vista parcial de las relaciones entre los conceptos, partiendo del dominio hacia el rango.

Tabla 5-Relación entre conceptos (Anexo 5)

Dominio	Relación	Rango
Análisis de competencias específicas	tiene	Actividades de aprendizaje seleccionadas
		Actividades de enseñanza seleccionadas
		Apoyos didácticos
		DCG instrumentales seleccionadas
		DCG interpersonales seleccionadas
		DCG sistemáticas seleccionadas
		Evaluación
		Fuentes de información
Horas teórico-prácticas		

Para su visualización gráfica, se diseñó un mapa conceptual mostrado en la **Figura 3**, donde se pueden observar las superclases de la ontología identificando de un color de relleno aquellas que contienen subclases.

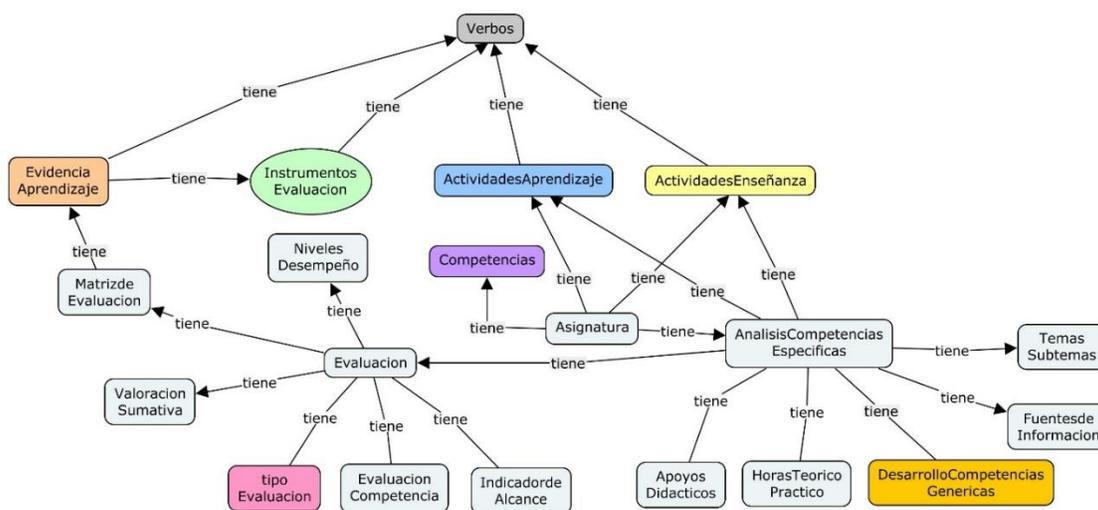


Figura 3-Diagrama de relaciones binarias

Una parte crucial de la ontología es la que abarca la Taxonomía de Bloom que está relacionada mediante los verbos de dominio a las clases correspondientes a las actividades de aprendizaje, actividades de enseñanza, instrumentos de evaluación y a las evidencias de aprendizaje, tal y como se muestra en la **Figura 4**.

Además, la subdivisión de la Taxonomía en sus 6 clases principales (análisis, aplicación, conocimiento, síntesis, comprensión y evaluación) adjunto a la correspondencia con las actividades, implica una extensa lista de relaciones que pueden inferirse, por lo que el uso de la ontología para almacenar la información de la taxonomía y la deducción del motor de inferencia para crear nuevas relaciones a partir de las ya existentes, son la principal justificación para su utilización en la presente investigación.

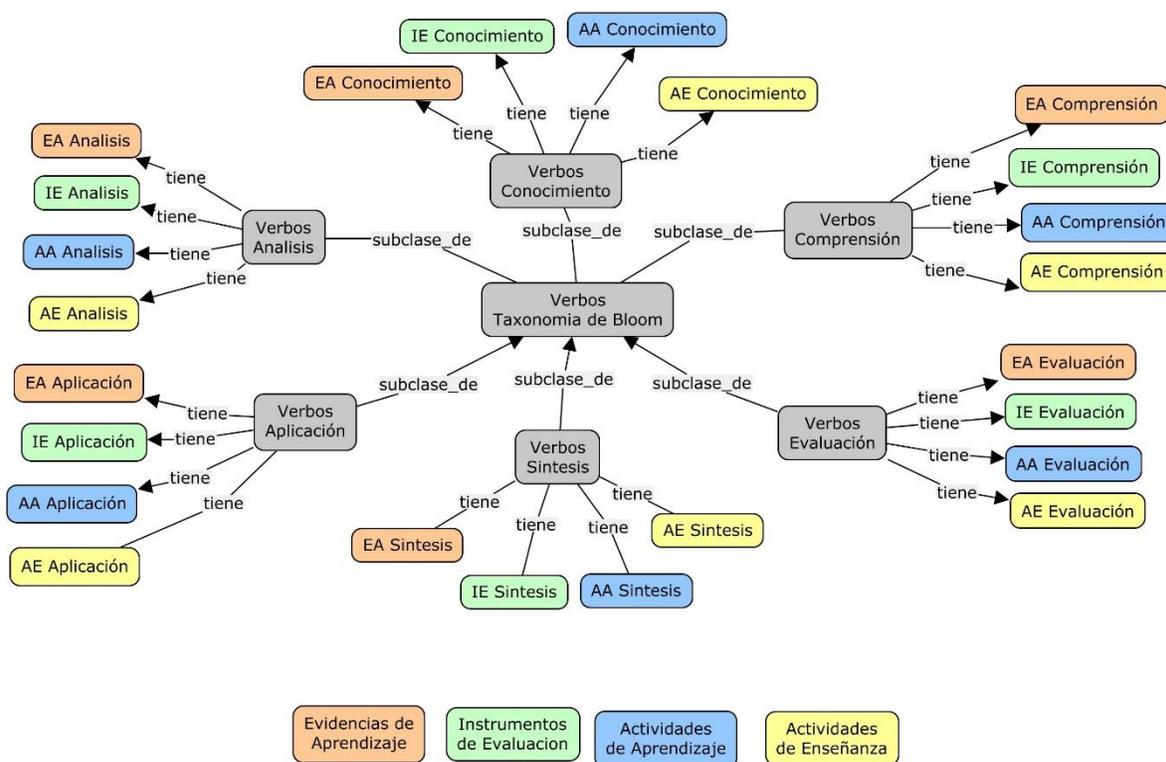


Figura 4-Vista de la Taxonomía de Bloom en la ontología

3.4.3- Procesos

La redacción de la ID por parte del docente se realiza directamente en la plataforma de capacitación en base a su experiencia y el temario de la asignatura, siendo la fuente de datos para el módulo de análisis que se encargará de la evaluación. Una vista genérica del proceso se muestra en la **Figura 5**. De la misma forma, en la **Figura 6** se visualiza más a detalle el proceso distinguiendo 3 formas principales:

- **Forma 0:** Contiene los datos de la ID que permitirán consultar la ontología personal. Los datos de la Forma 0 son: clave, año, semestre y RFC.
- **Forma 1:** Este formulario contiene los datos de la asignatura que son sacados en su mayor parte del temario de la asignatura. Entre los datos de la Forma 1 se encuentran la clave, nombre, caracterización, intención didáctica, carrera, aula, entre otros.
- **Forma 2:** Formulario referente a las unidades de la asignatura. En él, se va a redactar las competencias específicas a cumplir en la unidad y, a su vez, redactará las actividades de enseñanza, actividades de aprendizaje, evidencias de aprendizaje, instrumentos de evaluación y otros instrumentos didácticos que utilizará para cumplir y evaluar el nivel de alcance de la competencia específica planteada.

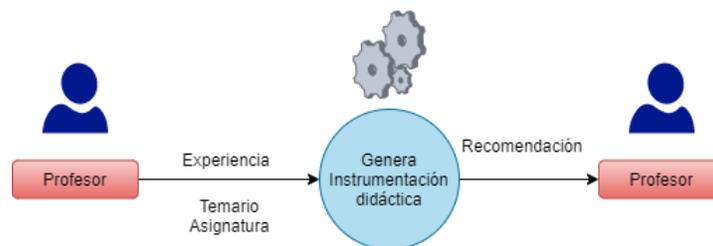


Figura 5-Diagrama de contexto (Nivel 0)

Para una visualización más precisa, en la **Figura 7** se desglosan 5 procesos de los cuáles los tres primeros corresponden a la población de la ontología y se pueden describir de la siguiente manera:

- **Proceso 0:** Se encarga de validar la existencia de la ontología propia del docente. Si no existe el registro, se crea con los datos de la Forma 0. Caso contrario, se continúa con proceso 1.

- Proceso 1:** En este apartado, se van a capturar los datos correspondientes a la Forma 1 (asignatura). A diferencia del proceso anterior, al hacer la validación de existencia de este registro y confirmar la existencia previa de la Forma 1, este será modificado, almacenando los nuevos datos. Si no se tiene registro, este será creado y posteriormente almacenado.
- Proceso 2:** Cuando se capturan los datos de la Forma 2 (unidad), es necesario no solo validar la existencia de la unidad si no el número al cual corresponde, debido a que una ID cuenta con varias unidades. Por eso, se hace una consulta a la ontología y, en caso de que exista un registro previo de la Forma 2, se procede a hacer la modificación y almacenar los nuevos datos. Caso contrario, si no se tiene registro se crea uno nuevo y se registra la Forma 2.

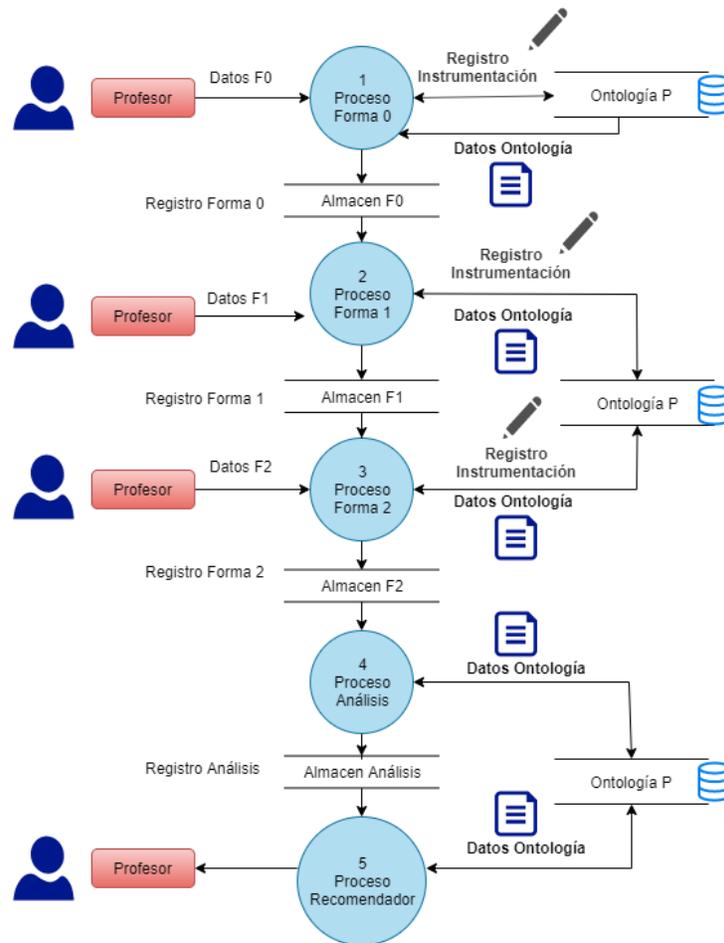


Figura 6-Diagrama de nivel inferior (Nivel 1)

Una vez que se ha agregado una unidad completamente, es posible analizarle, permitiendo seguir el proceso hacia el paso número 4, en el cual se utilizan técnicas de PLN y consultas a la ontología (con SPARQL). Con los resultados obtenidos, se procede al proceso 5 el cual es el encargado de otorgar la retroalimentación para terminar el ciclo de evaluación.

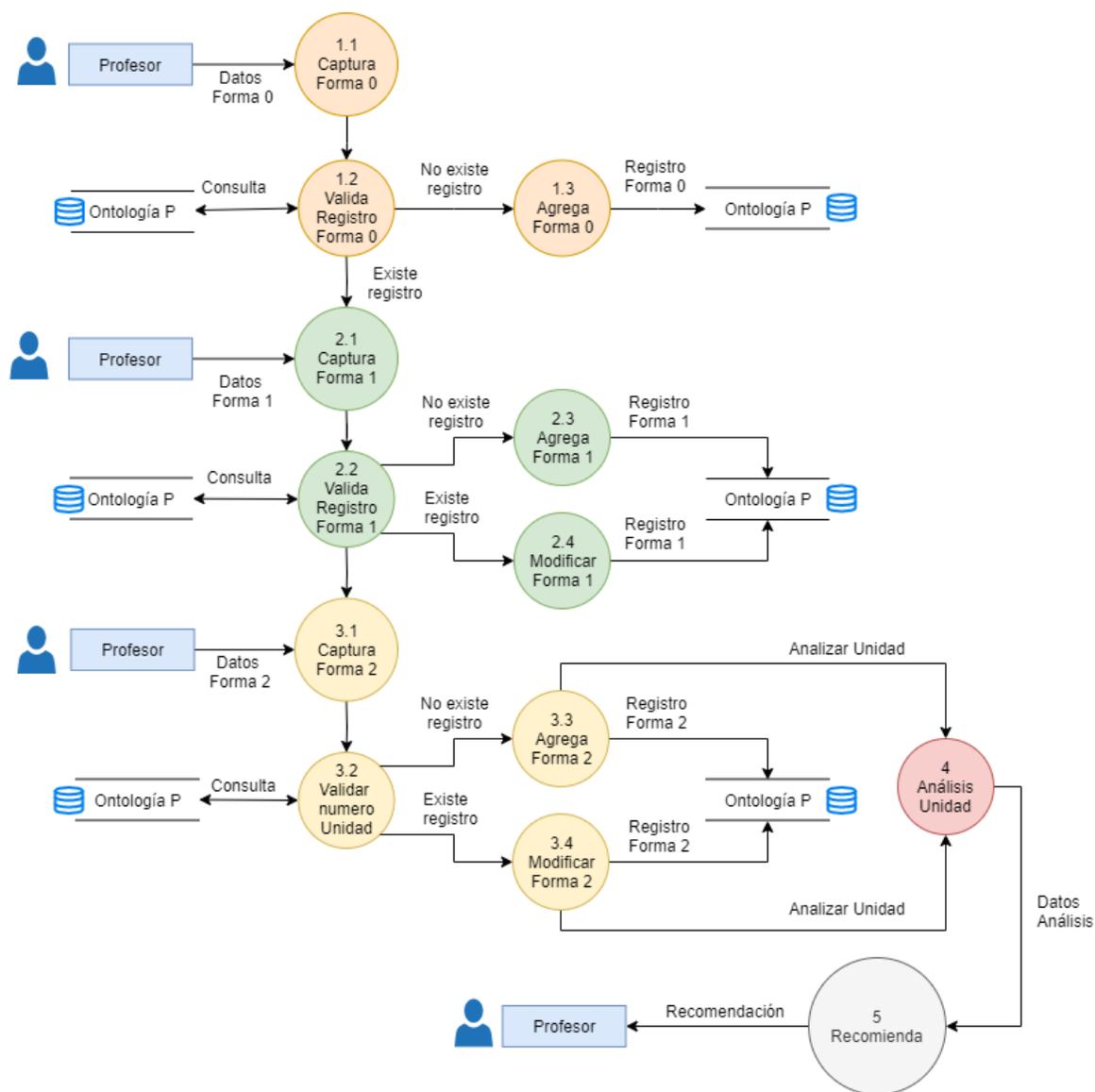


Figura 7-Diagrama de nivel superior (Nivel 2)

3.5- Diseño de algoritmo de PLN

En la **Figura 1** mostrada al inicio del capítulo con la arquitectura completa del sistema, se aprecia la composición del módulo de análisis en dos secciones: la primera corresponde a la población de la ontología con los datos recibidos de la plataforma de capacitación; mientras que la segunda es el proceso de análisis utilizando técnicas de PLN y consultas al modelo con SPARQL para la recuperación de información. Para una mejor visualización, en la **Figura 8** está plasmada la secuencia general del algoritmo de PLN dividido en 5 procesos:

- **Proceso 1:** Segmentación de la competencia específica utilizando un análisis morfológico para la detección de las partes esenciales (verbo, actividad y finalidad o condición de ejecución), así como la determinación del dominio al cual pertenece basándose en la Taxonomía de Bloom.
- **Proceso 2:** La condición de ejecución de la competencia debe de estar relacionada con al menos una de las actividades de aprendizaje que han sido redactadas. A la vez, estas deben de estar enfocadas en alguno de los temas de la unidad.
- **Proceso 3:** Análisis de los objetos de aprendizaje (actividades de aprendizaje, actividades de enseñanza, evidencias de aprendizaje e instrumentos de evaluación) que se han capturado en la plataforma. Para ello, se usa como base las actividades predeterminadas almacenadas en la ontología de acuerdo con el dominio de la competencia específica de la unidad simplificadas en el proceso anterior. En este paso, no importa el orden en el cual sean evaluados.
- **Proceso 4:** Para cada actividad que se plantee, es necesario tener una evidencia de dicho aprendizaje. en ese sentido, se realiza un análisis bidireccional de la relación entre las actividades de aprendizaje que se han redactado y las evidencias plasmadas en la ID.
- **Proceso 5:** Las evidencias que el alumno entrega deben de ser calificadas, siendo necesario que se elija aquel instrumento de evaluación que más se adecúe a la tarea. La ID contiene una matriz de evaluación, por lo que esta relación está ligada 1-1 a diferencia del proceso anterior donde no se conoce que actividad está relacionada con que evidencia.

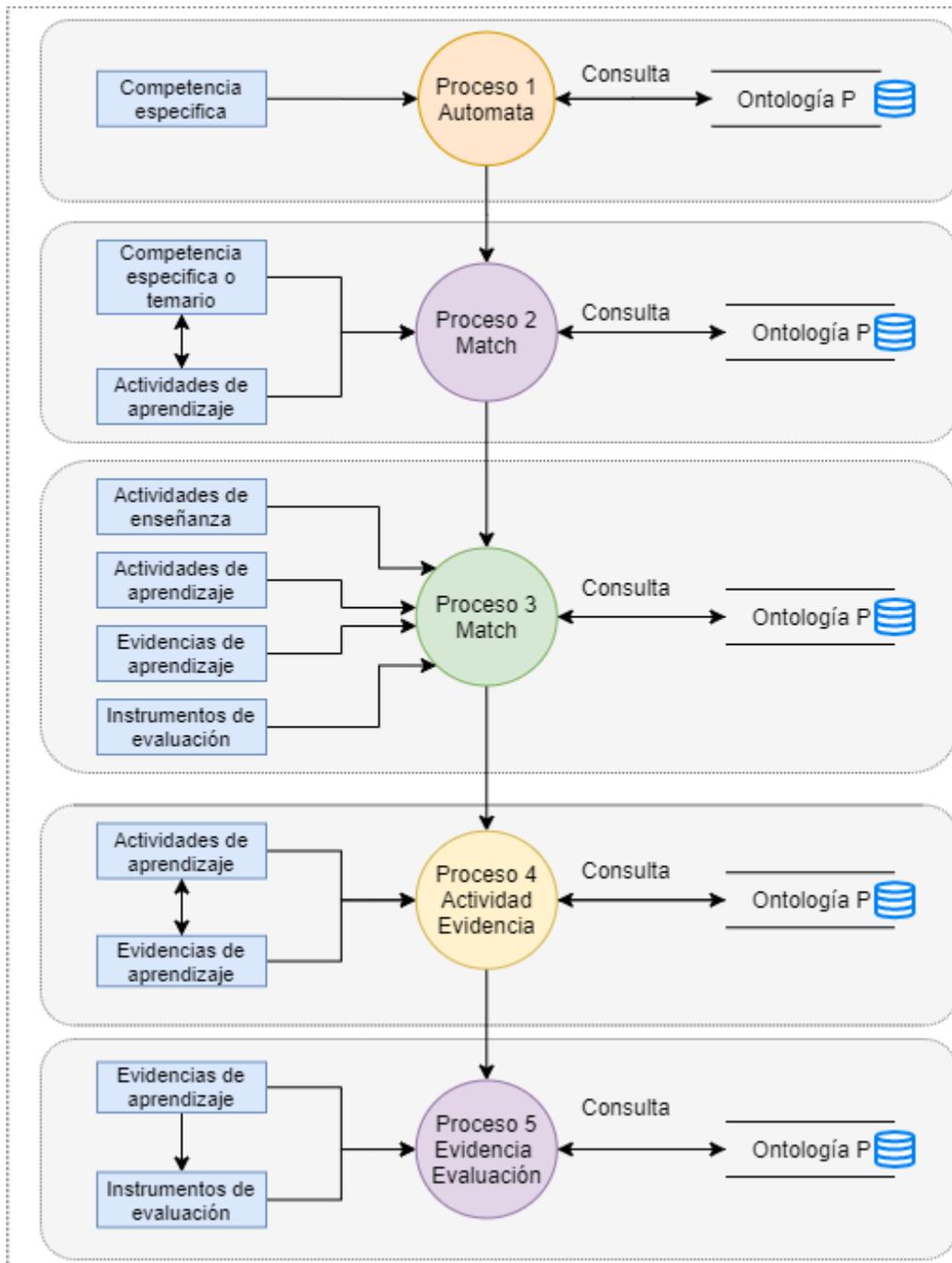


Figura 8-Secuencia de algoritmo de análisis de PLN

Como se puede observar, el proceso de evaluación es secuencial debido a lo cual es necesario completar cada proceso para poder continuar al siguiente. También, cada proceso otorga un resultado, por tal motivo es posible hacer el análisis por separado y no tener que esperar a tener toda la ID terminada para su evaluación.

3.5.1- Autómata

En México[41], las competencias específicas están redactadas en base a la Taxonomía de Bloom clasificada en seis niveles que contienen verbos claves para su identificación. En ella, los niveles van aumentando gradualmente de complejidad y, generalmente suelen ser representado como una pirámide, jerarquía que ayuda a clasificar tanto el verbo como a los objetos de aprendizaje. Partiendo de esa información y del conocimiento de los expertos, se puede concluir que una competencia está compuesta por tres secciones:

competencia = verbo + objeto + (finalidad y/o condición de ejecución)

Siguiendo ese patrón en su redacción, la detección de las entidades que componen la competencia se realiza mediante un análisis morfológico para determinar el papel gramatical que desempeña cada una de las palabras dentro de la oración. La **Figura 9** contiene el flujo del análisis partiendo desde el texto redactado por el profesor hasta la obtención de la lista de actividades en base al dominio de la taxonomía.

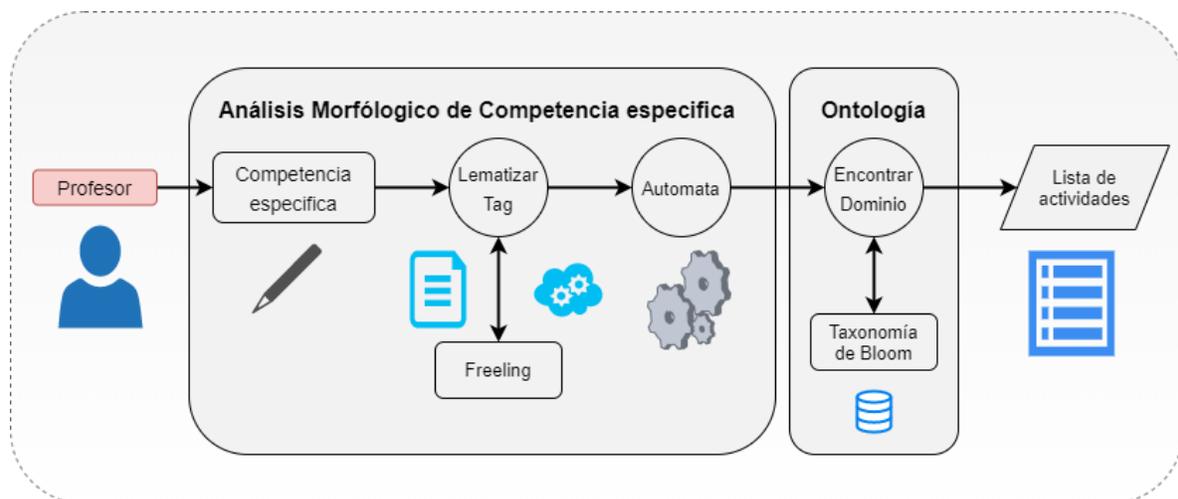


Figura 9-Secuencia de análisis competencia específica

Para el análisis morfológico, se emplea la librería de Freeling² que otorga el etiquetado de la categoría gramatical de las palabras. El autómata utiliza para la segmentación aquellas palabras cuya etiqueta gramatical³ indique que es un verbo, conjunción, sustantivo,

² <http://nlp.lsi.upc.edu/freeling/node/1>

³ <https://www.cs.upc.edu/~nlp/tools/parole-sp.html>

pronombre o inclusive signos de puntuación, cuyas variantes permiten realizar el cambio entre los estados del autómata, secuencia que se puede observar en la **Figura 10**, donde se muestra un estado inicial (S0) y tres posibles estados finales (S0, S3, S5). En cuanto a su alfabeto es el siguiente:

- **V:** El token es un verbo.
- **NP:** Sustantivo propio.
- **CC:** Conjunción coordinada.
- **PR:** Pronombre de tipo relativo.
- **CS:** Conjunción subordinada.
- **FC:** Signo de puntuación de tipo coma.
- **X-Y:** Cualquier etiqueta gramatical que sea excepción de las anteriores.
- **E1-E2-E3:** Representan excepciones a la regla.

Para determinar el dominio de una competencia, esta debe de comenzar con un verbo o un sustantivo propio, de no ser así no es posible determinar su dominio. Además, se toma en consideración que se puede tener más de un verbo de dominio en la competencia, siendo la transición del estado S1 al S2 en el cual se determina esta acción. En cuanto a las excepciones de la regla, son situaciones en las cuáles la redacción contiene elementos únicos que provocan que no se pueda hacer la segmentación de acuerdo con el esquema convencional, siendo estas excepciones quienes permiten realizar la separación y obtener la finalidad y/o condición de ejecución.

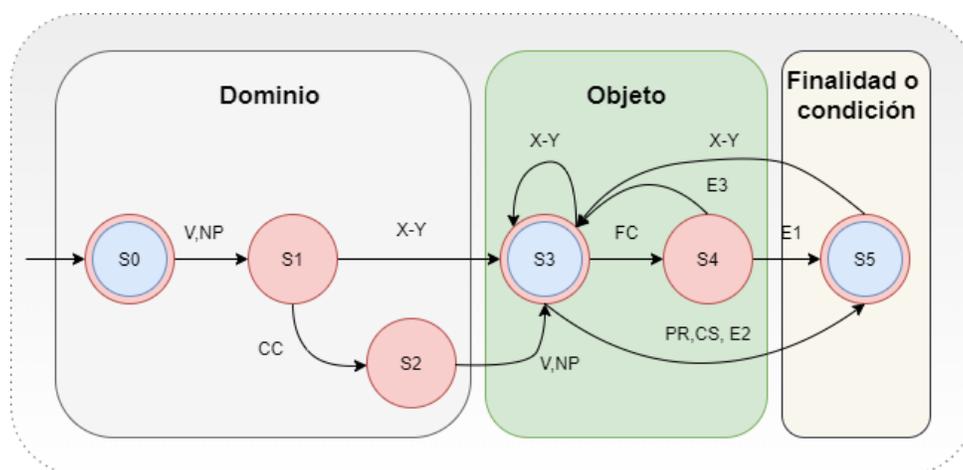


Figura 10-Autómata para segmentar competencia específica

Por ejemplo, si se analiza la competencia “*Comprende y aplica estructuras de datos lineales para solución de problemas*”, se consigue el etiquetado de la **Figura 11**, desde el cual se pueden obtener las siguientes transiciones:

- **V:** Pasa de S0 a S1.
- **CC:** Encuentra una conjunción. Pasa de S1 a S2.
- **V:** Encuentra verbo. Pasa de S2 a S3.
- **NC, SP, NC, AQ:** Forma parte de X, se mantiene en S3.
- **SP:** La palabra “*para*” indica una finalidad y forma parte de las excepciones de E2, por lo que se pasa de S3 a S5.
- **NC:** Regresa al estado S3 y se mantiene debido a que su etiqueta se considera como Y, por lo que el estado final es el S3.

Original: Comprende	Lematizada: comprender	Papel Estructural: VMM02S0
Original: y	Lematizada: y	Papel Estructural: CC
Original: aplica	Lematizada: aplicar	Papel Estructural: VMM02S0
Original: estructuras	Lematizada: estructura	Papel Estructural: NCFP000
Original: de	Lematizada: de	Papel Estructural: SP
Original: datos	Lematizada: dato	Papel Estructural: NCMP000
Original: lineales	Lematizada: lineal	Papel Estructural: AQ0CP00
Original: para	Lematizada: para	Papel Estructural: SP
Original: solución	Lematizada: solución	Papel Estructural: NCF5000
Original: de	Lematizada: de	Papel Estructural: SP
Original: problemas	Lematizada: problema	Papel Estructural: NCMP000

Figura 11-Ejemplo de análisis Morfológico de la competencia específica

Como resultado de la segmentación se obtiene lo siguiente:

- **Verbo:** Comprende y aplica
- **Objeto:** estructuras de datos
- **Finalidad:** para solución de problemas.

Con la detección de los verbos de la competencia, se determina su dominio haciendo una consulta a la ontología para recuperar las actividades y verbos del dominio utilizados en el proceso de match. Además, teniendo el tercer elemento identificado se busca su relación con las actividades de aprendizaje.

3.5.2-Composición de una ID

Una ID contiene información que está predefinida en los temarios de la materia donde se establecen los datos de la asignatura. La parte importante sobre la cual se centra la presente investigación es en el análisis por competencias específicas, donde a partir de la competencia y el temario de la unidad, se determina que actividades de aprendizaje y de enseñanza se utilizarán durante el curso, tal como la **Figura 12** que muestra un extracto de una ID.

Temas y subtemas para desarrollar la competencia específica (17)	Actividades de aprendizaje (18)	Actividades de enseñanza (19)	Desarrollo de competencias genéricas (20)	Horas teórico-prácticas (21)	Fecha programada del periodo		PAP 1OP (24)	PAP 2OP (25)	FD (26)	FJA (27)
					Inicial (22)	Final (23)				
1.1. Proceso Cognitivo Humano 1.2. Proceso de Aprendizaje Humano 1.3. Ergonomía cognitiva 1.4. Conceptos básicos 1.5. Fundamentos del diseño 1.6. Principios de diseño de objetos.	Mapa conceptual de términos básicos de la interacción usuario-computadora. Análisis de estudios de caso de interfaces de usuario. Ejercicios de aplicación de los principios de diseño al diseño de objetos virtuales.	Introducción a los conceptos y principios del diseño, mediante la elaboración de una presentación en Power Point y el uso de ejemplos. Presentación de estudios de caso. Formulación de preguntas.	Habilidades de investigación Trabajo en equipo	4-8	19/08	06/09				

Figura 12-Ejemplo de tabla que contiene los datos de una ID

Para complementar las actividades, se necesita que el alumno entregue evidencias de acuerdo con las actividades planteadas, por lo que lo anterior se complementa con una matriz de evaluación tal como en la **Figura 13**, donde se tiene una tabla con las evidencias seguido de los indicadores de alcance y los instrumentos elegidos para la evaluación. En el análisis, se considera solo las evidencias y los instrumentos, omitiendo el resto de la tabla.

Evidencia de aprendizaje (28)	%	Indicador de alcance (29)						Instrumentos de evaluación (30)
		A	B	C	D	E	F	
Mapa conceptual de términos. abordados.	40						X	Rúbrica
Reporte de estudio de caso.	30		X					Lista de control
Ejercicios de diseño de objetos virtuales.	30		X					Lista de control

Figura 13-Ejemplo de matriz de evaluación de una ID

3.5.3-Match de actividades de aprendizaje

Debido a que el Modelo EBC centra al alumno como el actor principal en el proceso educativo, las actividades de aprendizaje que se han elegido son un punto de inflexión al estar relacionadas directamente con la competencia de la unidad y el temario. Primeramente, el tercer elemento que compone a la competencia debe de estar relacionado directamente con al menos una de las actividades redactadas, de las que también se espera que esté complementada con algún subtema.

La **Figura 14** contiene un esquema base del proceso que se sigue para la búsqueda del match, considerando el preprocesamiento del texto y el ciclo repetitivo que se rompe cuando se ha encontrado una relación entre la competencia y al menos una de las actividades.

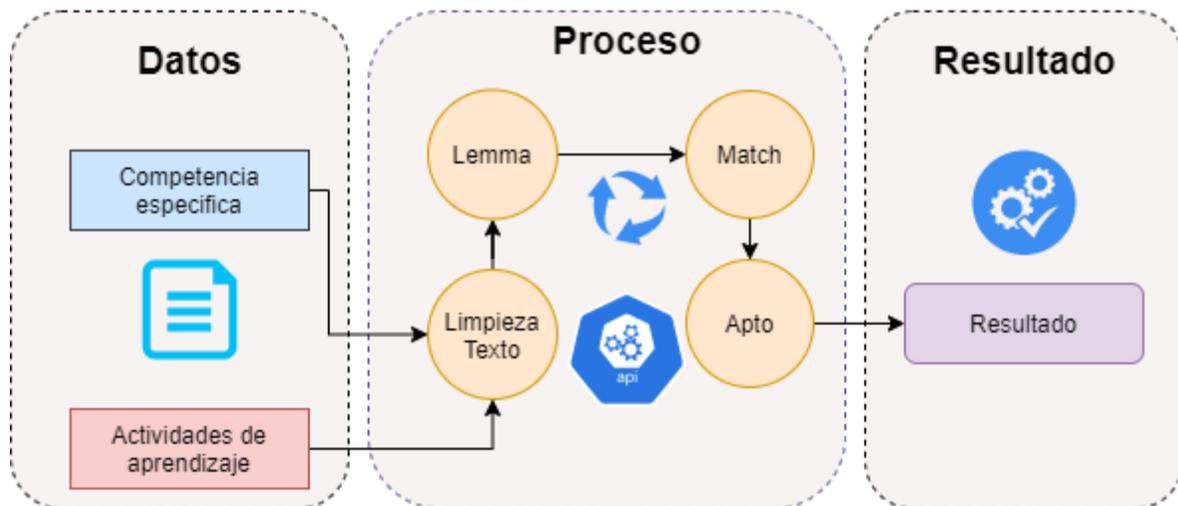


Figura 14-Proceso de match de actividades de aprendizaje

3.5.4-Proceso de Match

En el caso ideal, por cada subtema de una unidad se debe de tener una actividad de enseñanza, una actividad de aprendizaje y esta que genere una evidencia de aprendizaje que sea evaluada. Además, la redacción debe de tener la siguiente secuencia:

$$\textit{Actividad} = \textit{verbo} + \textit{actividad} + \textit{tema}$$

En dicha secuencia, el verbo debe de pertenecer al mismo dominio de la competencia específica al igual que la actividad elegida, por lo que aquí es donde se utilizan las listas resultantes del autómata. La **Figura 15** muestra el diagrama de flujo de los datos para la

extracción de las actividades redactadas. A la vez, se repite el proceso para los verbos que la acompañan para mantener el esquema.

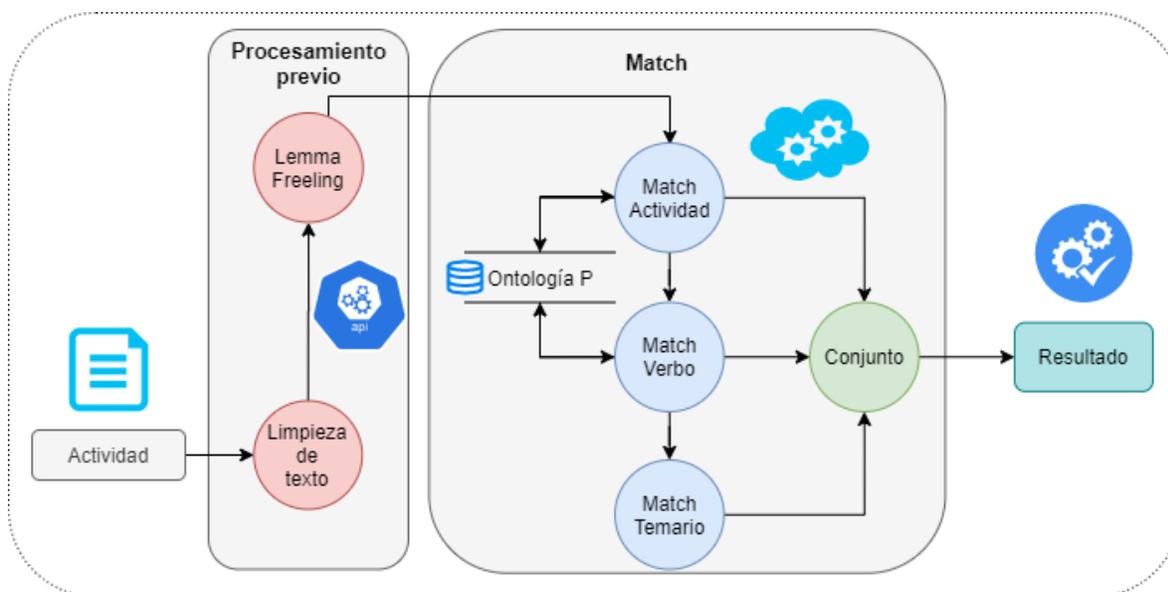


Figura 15-Revisión de redacción de las actividades utilizando Match

Se debe de tener en cuenta que las actividades de aprendizaje se consideran la estructura central de una ID al relacionarse directamente con la competencia específica, las actividades de enseñanza y con las evidencias de aprendizaje donde, a pesar de que se mantiene el patrón de búsqueda en la redacción, pueden tener o no un verbo que los acompañe y esto no influirá para determinar si es necesaria una revisión. Otra excepción se presenta en los instrumentos de evaluación, donde solo se busca la actividad omitiendo verbo y temario al carecer de ambos elementos.

3.5.3- Evidencia de actividades

Una vez que se ha redactado completamente una unidad de la asignatura es posible identificar si se ha encargado una evidencia por cada actividad que se ha planteado o viceversa. En la **Figura 16** se observa el diagrama sobre el cual se basará el algoritmo para encontrar la relación entre ambos elementos. Debido a que las actividades pueden resultar ambiguas y existen múltiples casos de excepciones a las reglas, se optará por realizar 3 tipos de análisis para obtener un único valor que estime si existe o no una relación. Los análisis por realizar en este paso son los siguientes:

- **Match:** La actividad debe de aparecer literalmente escrita en las evidencias, lo que ayudará a determinar si tiene alguna asociada. Para este punto, también se consideran las actividades similares que se pueden utilizar como posibles sinónimos, ampliando el catálogo de posibilidades haciendo uso de la ontología. Como resultado de este proceso se obtiene un valor booleano que indica la existencia o no de la relación.
- **Match difuso o parcial:** Un método alternativo es encontrar la correspondencia utilizando la distancia de Levenshtein, la cual mide la diferencia entre dos frases. En vísperas de que una *frase* debe de estar formada por un verbo, una actividad y un tema o finalidad, la evidencia tiende a contener estos elementos, por lo que puede ser tratado como una paráfrasis. En otras palabras, haciendo un análisis parcial de la *frase*, es posible encontrar un valor escalar que mida su similaridad y que otorgue un valor entre 0 y 100 donde 100 es una alta probabilidad de relación y 0 una nula.
- **Clúster:** El agrupamiento se da de manera no supervisada utilizando la técnica de vecino más cercano. Aun cuando la cantidad de datos es muy pequeña, es factible el algoritmo para encontrar patrones en el texto utilizando la técnica de extracción de características TFIDF. El resultado es un vector que contiene a la evidencia asociada de acuerdo con el análisis.

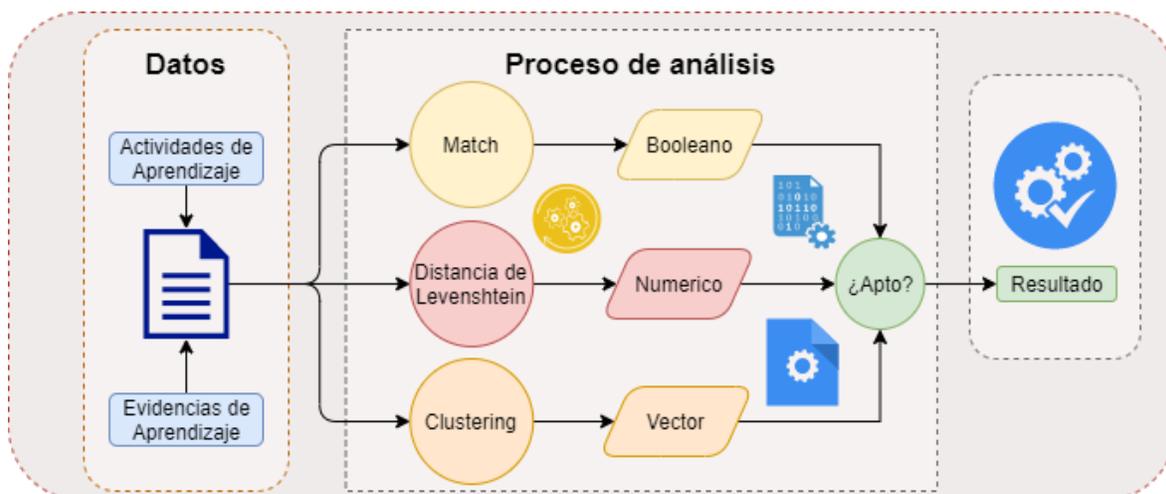


Figura 16-Relación entre actividades y evidencias de aprendizaje

3.5.4- Instrumentos de evaluación de evidencias

Finalmente, es necesario evaluar las evidencias de aprendizaje entregadas. La relación de ambos elementos se encuentra en la matriz de evaluación de la ID, lo que indica que es una correspondencia de 1:1 a diferencia del paso anterior donde se desconoce su posible relación. En el proceso, se analiza cada evidencia y se le asigna los instrumentos adecuados para su revisión en base al conocimiento de un experto en el área. Como resultado, la **Tabla 6** (ver **Anexo 6**) contiene un listado parcial de las evidencias, con los posibles instrumentos de evaluación.

Tabla 6-Relación entre Evidencias e instrumentos (ver **Anexo 6**)

Evidencia\Instrumento	R	L	E	P	X	G	A	G	H	I
actividad focal inductora	1	1	1	1	0	0	0	0	0	0
Analogía	1	1	1	1	0	0	0	0	0	0
Aprendizaje Basado en Análisis	1	0	0	1	0	0	0	0	0	0
Aprendizaje Basado en Problemas	1	1	0	1	0	0	1	0	0	0
Aprendizaje Basado en Proyectos	1	1	0	1	1	0	0	0	0	0
artículo revista	1	1	1	1	0	0	0	0	0	0
autoevaluación/heteroevaluación	1	1	1	1	0	0	1	0	0	0
búsqueda definición	1	1	1	1	0	0	0	0	0	0

Capítulo IV-Implementación

4.1-Introducción

En el presente capítulo, se estipula la implementación de la ontología y el algoritmo de Procesamiento de Lenguaje Natural (PLN) en base al análisis realizado en el capítulo anterior. Siguiendo la metodología Methontology, se procedió a la construcción del modelo utilizando las tablas con los conceptos e individuos del dominio, así como las relaciones representadas en los diagramas.

Después de su implementación, se analizaron los pasos seguidos para el algoritmo de PLN, partiendo de la codificación del autómata para la segmentación de la competencia específica de la Instrumentación Didáctica (ID). Así mismo, se muestra el pseudocódigo de algunas secciones para la detección de los patrones en el texto.

Mas adelante, se detallaron las consultas a la ontología utilizando el lenguaje SPARQL para la recuperación de la información y se contempló el uso del motor de inferencia para deducir nuevo conocimiento. Por último, se indicaron los valores que son enviados al módulo de recomendación para que realice su función de retroalimentar al docente en base a los resultados del módulo de análisis.

4.2- Implementación de la Ontología

Haciendo uso de Protégé en su versión 5.5, se implementó el modelo computacional diseñado en el capítulo anterior. Entre las principales bondades respecto a otro software de su tipo se encuentra lo intuitiva que es su interfaz, la facilidad en la captura de datos, la amplia cantidad de complementos que facilitan las tareas de construcción, anexo a la facilidad para la importación y exportación del modelo resultante en diferentes formatos. También, es posible utilizar complementos para la visualización gráfica de las relaciones y el hacer consultas con SPARQL.

Ahora bien, para el uso de este software fue necesario hacer unos ajustes en los nombres que define la metodología Methontology tal como definir los conceptos como clases dentro del software, o el uso de las propiedades de objeto para construir las relaciones entre los conceptos. En cuanto a los atributos tanto de clase como de instancia, el software los maneja como propiedades de dato, mientras que las instancias tienen por nombre individuos dentro de Protégé.

4.2.1.- Construcción de las clases

La codificación de la ontología está basada en la metodología Methontology donde se formalizó el conocimiento haciendo uso de los conceptos del dominio, las relaciones y los individuos que le componen. Primeramente, se jerarquizó las clases y subclases formando la taxonomía de esta donde se obtuvieron 21 clases principales mostradas en la **Figura 17** como una vista parcial de los 83 conceptos.

También, se puede observar a las actividades de aprendizaje separadas en predefinidas y seleccionadas. En la primera categoría se tienen subclases correspondientes a los dominios de la taxonomía de Bloom, separación que facilita su recuperación específica. La misma lógica se sigue para las actividades de enseñanza, las evidencias, los instrumentos y los verbos de la competencia.

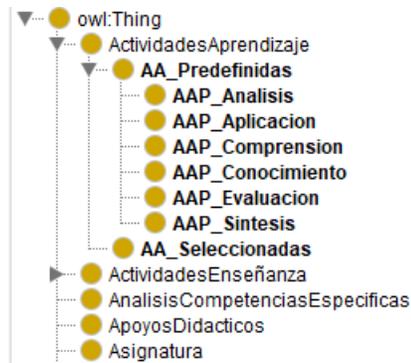


Figura 17-Vista parcial de los conceptos de la ontología

4.2.2- Relaciones entre conceptos

En Protégé, las relaciones entre los conceptos reciben el nombre de propiedades de objetos y permiten hacer inferencia de conocimiento mejorando la recuperación de la información al aportar por sí mismas conocimiento. En este modelo se cuenta con 49 propiedades de objeto diferentes.

La **Figura 18** contiene una vista parcial de las relaciones de la clase *análisis de competencias específicas* que funge como dominio, con otras clases que determinan su rango tales como *apoyos didácticos*, *actividades de aprendizaje*, *actividades de enseñanza*, *competencias instrumentales*, entre otras. Además del dominio y rango, Protégé admite establecer otras propiedades como declarar clases equivalentes, subpropiedades o inversos entre las clases.

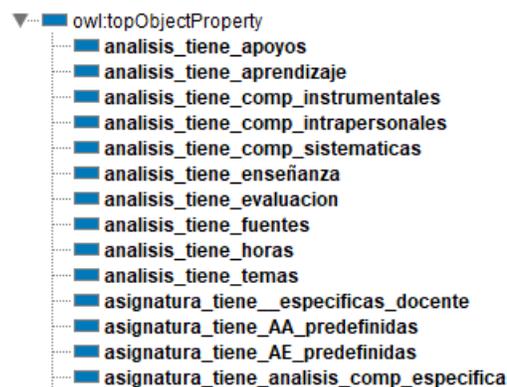


Figura 18-Vista parcial de las relaciones entre las clases

4.2.3- Propiedades de datos

En Protégé, los atributos que definen a las clases son llamados propiedades de datos. En este apartado se definieron las propiedades propias de las clases que permiten hacer la diferenciación entre ellas considerando para su formulación que los atributos son heredados por todas las subclases.

En la columna de la izquierda de la **Figura 19** se observan los atributos de las clases formuladas para este modelo que contiene un total de 33 diferentes. Una de las principales ventajas que presenta su declaración es que, si se formulan adecuadamente, pueden llegar a evitar la ambigüedad al momento de realizar la inferencia. Por otra parte, es necesario determinar el tipo de valor que será almacenado, tal como puede ser de tipo texto, número, flotante, booleano o inclusive fechas.



Figura 19-Vista parcial de los atributos de las clases

4.2.4- Instancias

Las instancias son representaciones de objetos en el dominio sobre el cual se está trabajando y su relación con las clases generalmente está dada del tipo “*esUn*”. La taxonomía de Bloom establece una lista predeterminada de verbos y actividades, por lo que es necesario hacer su declaración como instancias y así, estos puedan ser consultados por el módulo de análisis.

Como se puede observar en la **Figura 20**, el verbo seleccionado *indicar* se declara de tipo *verbo de análisis*, *verbo de aplicación* y *verbo de comprensión*, por lo que pertenece a dichas clases. Una de las opciones del software de Protégé es “*Same Individual As*”, la cual admite declarar individuos como iguales ayudando a inferir nuevo conocimiento. Por ejemplo, en la instancia “*ilustración*”, no es necesario declarar que es un verbo de *análisis*, *aplicación* y *comprensión*, al expresar que es igual que “*ilustrar*”, el motor de inferencia permite deducir que ambos pertenecen a las mismas clases.

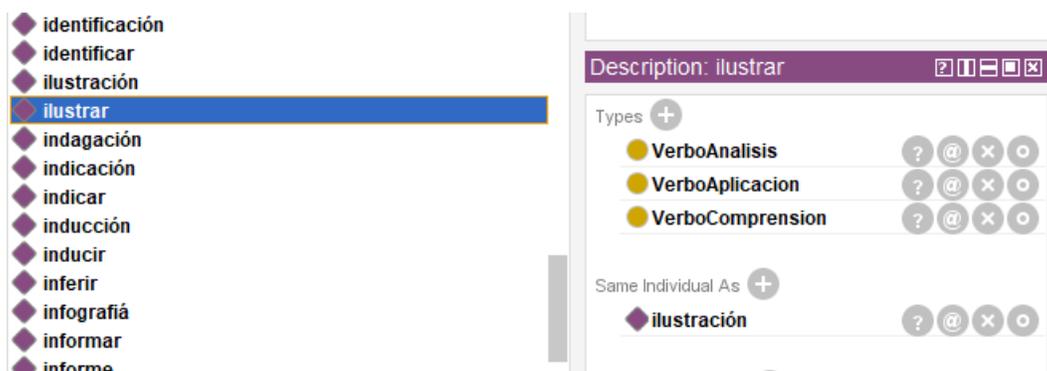


Figura 20-Vista parcial de las instancias

En total, se cuenta con 486 individuos que en su mayoría pertenecen a los verbos de la taxonomía y a las actividades predeterminadas. Para una mejor visualización de los pertenecientes a una clase, al seleccionarla en Protégé se enlistan sus individuos tal y como se muestra en la **Figura 21**, donde se visualizan algunos datos relevantes como lo son el tipo de dato o sus propiedades.

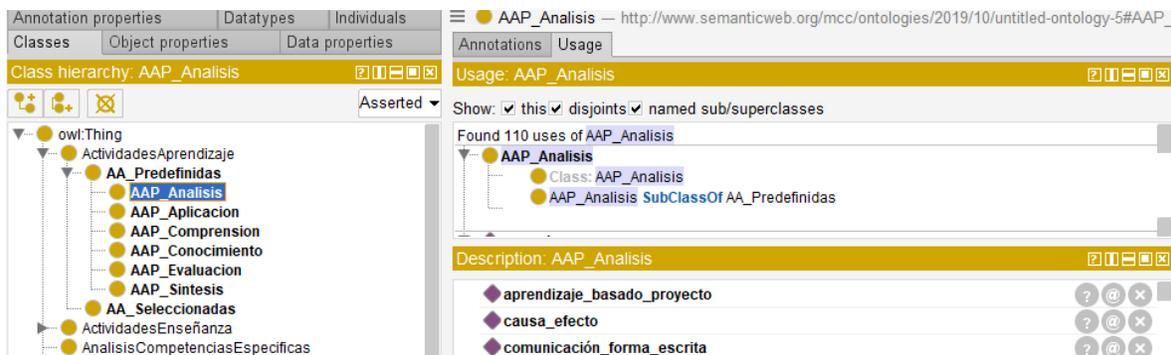


Figura 21-Visualización de las clases y sus individuos

4.2.5- Motor de inferencia

Como se explicó en el ejemplo anterior, el motor de inferencia permite la creación de nuevas relaciones a partir del conocimiento ya existente. Siguiendo con el verbo “*ilustrar*” que contenía como mismo individuo a “*ilustración*”, en la **Figura 22** se pueden observar con un contraste amarillo aquellas propiedades que han sido inferidas. Por ejemplo, se encontró que además de “*ilustración*”, también es equivalente a *bosquejo, dibujar, dibujo y fotografía*. Por consiguiente, se añadió la pertenencia a las clases *de actividades de aprendizaje, evidencias de aprendizaje y los instrumentos de evaluación* con los cuáles es posible evaluarle como evidencia. En ese sentido, el punto fuerte del uso de ontologías radica en esa capacidad de inferencia aun cuando el dominio es muy específico.

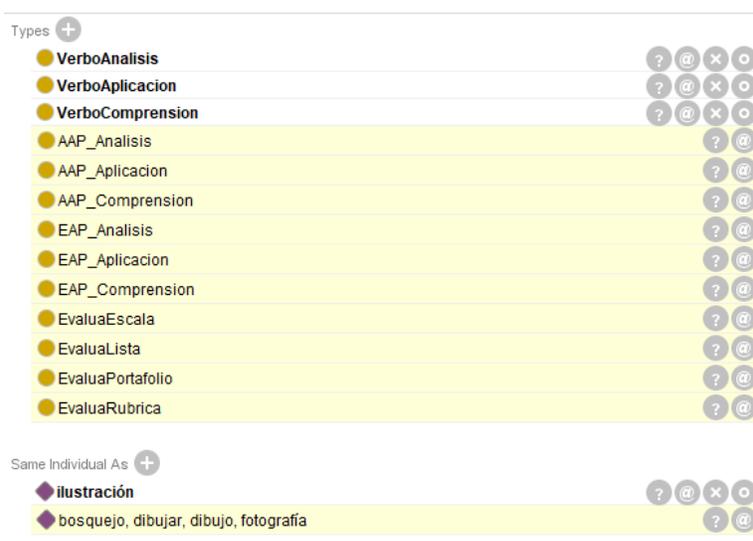


Figura 22-Deducción obtenida utilizando el motor de inferencia

4.2.6- Visualización gráfica de la ontología

El Software Protégé cuenta con complementos para visualizar gráficamente la ontología entre los que destaca OntoGraf que permite observar las relaciones entre las clases y subclases, así como los individuos que las componen, tal como lo muestra la **Figura 23** que contiene las relaciones del concepto de actividades de aprendizaje. Como se explicó anteriormente, se tiene un listado de actividades predefinidas, por lo que la mayor cantidad de relaciones se encuentra en este apartado clasificadas de acuerdo con la taxonomía y a su vez con los verbos de la categoría a la cual pertenecen.

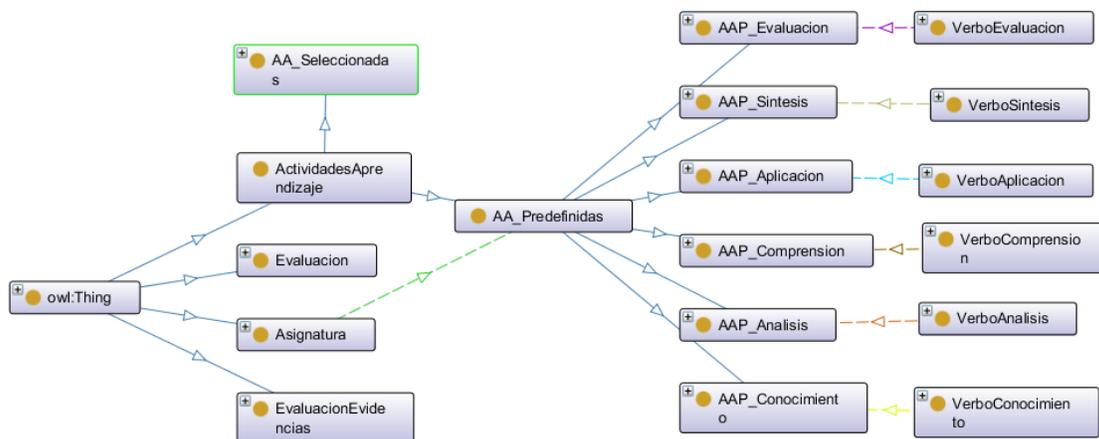


Figura 23-Visualización de las relaciones utilizando OntoGraf

4.3- Algoritmo de PLN

Si bien, una ontología admite organizar la información de un dominio en específico para facilitar su recuperación dada una consulta, es necesario el uso de otras herramientas que permitan analizar el texto en formato libre para la extracción de características, por lo que se desarrolló un algoritmo utilizando técnicas de Procesamiento de Lenguaje Natural (PLN) para cumplir con esa tarea.

La **Figura 8** (perteneciente al capítulo anterior) indica la división del algoritmo en cinco procesos principales partiendo desde el análisis de la competencia específica para determinar su dominio y relación con las actividades de aprendizaje. A su vez, estas se enlazan con las evidencias entregadas por el alumno que son evaluadas haciendo uso de instrumentos. Este proceso descrito brevemente es reiterativo por cada unidad, pero es secuencial dentro de la misma.

4.3.1.-Requisitos previos

El algoritmo está desarrollado en Python, lenguaje que contiene una extensa cantidad de librerías orientadas al ámbito de la Inteligencia Artificial (IA), el PLN y el Aprendizaje Máquina (mejor conocido como Machine Learning). Las librerías utilizadas son las siguientes:

- **re:** Permite trabajar con expresiones regulares con el fin de crear secuencias de caracteres para formar un patrón de búsqueda.
- **requests:** Se utiliza para enviar solicitudes HTTP a un servicio web. Su función en el algoritmo es realizar consultas a una API alojada en un servidor externo.
- **Freeling:** Librería de PLN cuya función principal es analizar a nivel morfológico (etiqueta POS) y sintáctico (Parsing) una oración con el fin de obtener la forma base (lemma) de las palabras y su papel estructural dentro de la misma.
- **NLTK:** Al igual que Freeling, NLTK es una librería de PLN con una amplia variedad de funcionalidades. En el algoritmo, ayuda en la eliminación de las palabras vacías y en la creación de n-grams para el proceso de Match.
- **Fuzzywuzzy:** Librería de lógica difusa cuya finalidad es encontrar la similaridad entre palabras utilizando la Distancia de Levenshtein, lo cual se puede interpretar como ¿Qué tan similares son dos frases entre sí?
- **Sklearn:** Orientada al Aprendizaje Máquina. Permite la extracción de características del texto y la agrupación (Clustering) de los mismos en base a un análisis estadístico que toma como base a las características similares de los objetos analizados para la formación de grupos de manera automática.
- **Rdflib:** Admite la comunicación con ontologías en formato OWL o RDF y realizar consultas de SPARQL desde Python.
- **Owrl:** Al igual que Rdflib, admite trabajar con modelos de conocimiento con la ventaja de que tiene un razonador deductivo entre sus múltiples funcionalidades.

Aparte de la importación de librerías, fue necesario la declaración de clases con el objetivo de almacenar en ellas los datos recibidos de la plataforma, las consultas al modelo y los resultados del análisis tanto del match entre las actividades, como la relación entre los elementos de la ID.

4.3.2- Autómata para segmentación de competencia

Cada unidad desarrollada en una ID parte de una competencia específica que indica cual es el objetivo que se desea alcanzar al terminar dicha unidad y, a partir de ella y el temario, se redactan las actividades a utilizar en el proceso de enseñanza-aprendizaje. Para segmentar la competencia, se diseñó un autómata para la detección de las entidades que le

componen, partiendo de un análisis morfológico con el que se determina el papel gramatical que desempeña cada una de las palabras dentro de la oración. La **Figura 24** muestra el pseudocódigo base del autómata implementado en Python, donde se utilizó la competencia como un vector de palabras para encontrar las posiciones de separación, obteniendo como resultado lo siguiente:

- **Verbo:** Competencia [: Posición 1].
- **Objeto:** Competencia [Posición 1: Posición 2]
- **Finalidad o condición 1:** Competencia [Posición 2: Posición 3]
- **Finalidad o condición 2:** Competencia [Posición 3:]

```
Por pos, token en Enumeración competencia:
Si token["tag"] inicia con ('V','NP'):
    Si competencia[SIGUIENTE TOKEN]["tag"] inicia con ('CC'):
        Posición1 = 3
    Si no: Posición1 = 1
    Si token["lemma"] esta en Ontología: dominio = True

Si token['tag'] inicia con ('PR','CS',E2,FC):
    Si E3: continuar
    Si no:
        Si posición2 Igual 0: Posición2 = pos
        Si no: Posición3 = pos
```

Figura 24-Pseudocódigo base del autómata

Debido a que una competencia puede tener más de un verbo que define su dominio, la posición 1 puede variar su valor entre 1 y 3. Además, es vital que el verbo forme parte de la ontología, por lo que se realiza una consulta para corroborarlo. En algunos ejemplos, se encontró que el docente no utilizaba la redacción de los verbos en el tiempo indicado, tal es el caso de “*elaboración*” donde el papel gramatical que se le otorga es el de un sustantivo propio aun cuando se deriva del verbo *elaborar*, motivo de la integración de NP para la identificación del dominio. También, una competencia puede tener solo una finalidad o condición de ejecución. En esos casos, la finalidad 1 parte de la posición 2 hacia el final de la competencia, dejando la finalidad 2 vacía.

En el análisis morfológico, se utilizó la librería Freeling para realizar el etiquetado y obtener la forma base de la palabra. Por ejemplo, la **Figura 25** muestra la segmentación de la competencia utilizando el autómata, lo que dio como resultado lo siguiente:

- **Verbo:** Diseña y desarrolla
- **Objeto:** Una aplicación móvil
- **Finalidad o condición:** A partir de un módulo de negocio seleccionado.

```

"token": Diseña      "lemma": diseñar    "tag": VMIP350
"token": y           "lemma": y          "tag": CC
"token": desarrolla  "lemma": desarrollar "tag": VMIP350
"token": una         "lemma": uno        "tag": DIØFSØ
"token": aplicación  "lemma": aplicación "tag": NCF5000
"token": móvil       "lemma": móvil      "tag": AQØCSØØ
"token": a_partir_de "lemma": a_partir_de "tag": SP
"token": un          "lemma": uno        "tag": DIØMSØ
"token": modelo      "lemma": modelo     "tag": NCCSØØØ
"token": de          "lemma": de         "tag": SP
"token": negocio     "lemma": negocio    "tag": NCMSØØØ
"token": seleccionado "lemma": seleccionar "tag": VMPØØSM

```

Figura 25-Resultado de análisis morfológico

4.3.3- Consultar Ontología

Con la competencia segmentada y los verbos que la componen identificados, se procedió a determinar el dominio. Con el fin de realizar esta tarea, se creó una función para realizar las consultas sin la necesidad de estar reescribiendo el código siendo necesario solo enviar la consulta y el archivo del modelo en formato OWL como parámetros de la función tal y como se visualiza en la **Figura 26**. Utilizando la función, la consulta para la obtención del dominio de una actividad es la siguiente:

$$\text{dominio} = \text{"SELECT * WHERE \{onto: verboC rdf:type ?m\}}$$

Mientras tanto, para la lista de verbos del dominio, se utiliza el resultado de la consulta anterior de la siguiente manera:

$$\text{predeterminado} = \text{"SELECT * WHERE \{?m rdf:type onto: dominio\}}$$

```

Función ConsultarOntologia(consulta, ontologia):
  Predeterminado = ""PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
  PREFIX owl: <http://www.w3.org/2002/07/owl#>
  PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
  PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
  PREFIX onto: <http://www.semanticweb.org/mcc/ontologies/2019/10/untitled-ontology-5#>""
  consultar = ontologia.Query(predeterminado + consulta)
Por encontrado en Consultar:
  Agregar a resultado(encontrado[0].Separador(#)[1])
Regresar resultado

```

Figura 26-Función para consultas a la ontología

El proceso anterior se repite para la recuperación de las actividades en base a su dominio y, en dado caso de que se recupere más de uno, en la consulta se utiliza *UNION* para recuperar todos los elementos pertenecientes a cada uno de ellos.

4.3.4- Match

El trabajo previo realizado forma parte del preprocesamiento de los datos de la unidad, solo contemplando la competencia específica como texto por parte del docente. Para facilitar la interpretación de este proceso, cada una de las actividades que conforman el listado de actividades de aprendizaje, actividades de enseñanza, evidencias de aprendizaje e instrumentos de evaluación será llamado *frase* en el resto del capítulo.

La **Figura 27** muestra el pseudocódigo base del algoritmo de match, el cual comienza con un preprocesamiento de las *frases* con la finalidad de eliminar caracteres especiales y conectores que no aportan información relevante al análisis. En esta tarea, se utilizaron expresiones regulares y NLTK para la limpieza del texto en conjunto con Freeling para obtener el lemma (raíz) de las palabras y su papel gramatical.

Con el texto resultante, se evalúa si el elemento que se está analizando es una actividad de aprendizaje, de ser verdadero se procede a buscar el match con la competencia específica utilizando una función que calcula la Distancia de Levenshtein entre la

competencia y cada una de las actividades; de encontrarse relación, se retorna la actividad y un valor booleano de verdadero.

A continuación, se procede a realizar el análisis de las *frases* individualmente utilizando un ciclo en el proceso. En primera instancia, se utiliza NLTK para la obtención de n-grams a partir del texto lematizado, con la intención de que las actividades que contengan dos o más palabras puedan ser evaluadas en conjunto como es el caso de *mapa conceptual*. Así pues, se busca el match entre las actividades recuperadas previamente de la ontología y la *frase* actual utilizando la Distancia de Levenshtein ante la posibilidad de existencia de errores ortográficos por el texto en formato libre. También, es posible que una actividad pueda repetirse dentro de una misma *frase* debido al uso de n-grams. Para evitarlo, por cada nueva actividad que se ha encontrado en la *frase*, se realiza una consulta al modelo para recuperar la lista de similares y, en dado caso de que ya exista, no es anexada al resultado.

En cuanto a los verbos, no es necesario hacer el análisis a toda la *frase*, por lo que solo se realiza en aquellos tokens cuyo papel gramatical indique que son un verbo o un sustantivo propio. La misma lógica se sigue para encontrar el tema al cual pertenece la *frase*, donde es comparada contra el temario de la unidad para comprobar que se ha utilizado alguno de los temas y así, verificar que el docente ha cumplido con los 3 elementos de la *frase*: verbo, actividad y tema.

```

texto_lematizado texto_original = preprocesamiento_texto(frases)
Si evaluarAA es igual a True:
    actividad_aprendizaje_competencia(texto_lematizado)
Por frase en texto_lematizado:
    n-grams = ObtenerNGrams(frase)

    Por actividad en lista_actividades:
        Distancia = Distancia_Levenshtein(actividad,palabra)
        Si distancia es mayor a 90:
            actEncontrada = True
            actividad_enc = actividad
            QuitarSimilarOntologia(actividad)
            Break

    Por palabra en frase:
        Si palabra["tag"] inicia con ("V","NP"):
            Por verbo en dominio:
                Valor = Distancia_Levenshtein(palabra["lemma"],verbo)
                Si valor es mayor a 90:
                    Conjunto = verbo + actividad_enc
                    valor_conjunto = Distancia_Levenshtein(conjunto,frase)
                    Si valor_conjunto es mayor a 90:
                        verbo_encontrado = True
                        verbo_enc = verbo
            buscarActividadTemario(temario_lematizado,frase)

```

Figura 27-Pseudocódigo base para obtener el Match

Para terminar con este proceso, se retornan los siguientes valores:

- **Revisar:** Indica si es necesario o no que se realice una revisión de la *frase*.
- **Original:** *Frase* original analizada.
- **Lematizada:** *Frase* pre procesada.
- **Actividad:** Su valor corresponde a la actividad con la que se ha hecho match. Si no se hizo ningún Match, su valor es "NA".
- **Verbo:** Contiene el verbo que acompaña a la actividad.
- **Tema:** Es el tema al cual pertenece la frase analizada.

Los resultados anteriores son agregados a un arreglo en formato JSON que es enviado al módulo de recomendación. Cabe resaltar que este proceso analiza por separado las actividades de aprendizaje, actividades de enseñanza y evidencias de aprendizaje. Como omisión, el análisis es parcial para el caso de los instrumentos de evaluación, debido a que solo se identifica la actividad al carecer en su redacción de un verbo y tema asociado.

4.2.5.-Relación entre Actividades y Evidencias de Aprendizaje

Para cada actividad que el docente plantee para sus alumnos, es necesario que estos entreguen una evidencia, de ahí la importancia de relacionarles entre sí. No obstante, las posibilidades de las relaciones que se pueden dar en este apartado son numerosas. En consecuencia, fue necesario recurrir a 3 métodos diferentes (match, Distancia de Levenshtein y Clustering) para en conjunto obtener un único resultado que les relacione. En el caso ideal, debe existir una evidencia por cada actividad bien definida, pero se presentan casos en que estas no tiene asociada una evidencia o viceversa, siendo necesario tomar estas excepciones para que el análisis fuese bidireccional.

En el pseudocódigo de la **Figura 28** se representa el análisis base para las correlaciones. Partiendo del proceso de match, se busca la existencia literal de la actividad encontrada en el proceso anterior en ambos grupos de elementos, obteniendo como resultado un valor booleano para indicar si ha sido encontrado. Aquí son utilizados en conjunto aquellos elementos similares que se han inferido desde el modelo con el fin de aumentar el rango de posibilidades.

Por otra parte, debido a que la redacción entre la *frase* de actividades y evidencias tienen rasgos similares, estos pueden ser considerados como una paráfrasis de su contraparte, por lo que se otorga un umbral de similitud entre ambos para medir su correlación. También, considerando que el proceso de análisis es bidireccional, la función es llamada dos veces cambiando únicamente el orden en que se presentan las listas. Este factor es importante cuando la cantidad de evidencias es diferente al de actividades o, cuando algún elemento de uno de los grupos no tiene un semejante en el otro.

Posteriormente, el tercer análisis es el agrupamiento utilizando Sklearn para la extracción de características mediante la técnica TF-IDF que permitió hacer la clasificación bajo la técnica de vecino más cercano. A pesar de ser una técnica no supervisada, el resultado obtenido es favorable debido a que la actividad y la evidencia comparten particularidades que permiten diferenciarles del resto del listado. Como resultado de este agrupamiento, se obtuvo un vector con las actividades agrupadas.

```
Funcion relación Distancia Match(lista1, lista2):  
  Por valor1 en lista1:  
    Por valor2 en lista2:  
      Si valor1['actividad'] es igual a valor2['actividad']:  
        encontradoMatch = True  
        resultado = Distancia Levenshtein (valor1['frase'], valor2['frase'])  
      Si resultado mayor o igual a 70:  
        encontradoL = True  
  relación Distancia Match (actividades, evidencias)  
  relación Distancia Match (evidencias, actividades)  
  
  vector = ObtenerTFIDF(actividades, evidencias)  
  matriz = obtenerMatrizSimilaridad(vector)  
  p = afinarPropagacion(matriz)  
  ClasificarClustering(p)
```

Figura 28-Seudocódigo base del análisis de relaciones

Finalmente, fue necesario unificar los resultados para obtener un único valor que represente el grado de relación. De los tres procesos, se le dio más relevancia al de match, al considerarse necesario que la actividad aparezca en ambos textos. Después, se optó por la *frase* con mayor similitud también estuviese en el vector de agrupamiento. Finalmente, se utilizaron métricas para evaluar posibles combinaciones entre los tres resultados obteniendo

un umbral entre aquellas que son aceptables, necesitan revisión o no se dispone de algún elemento que le correlacione. Como consecuencia, el JSON resultante de este paso contiene la siguiente información:

- **Actividad:** Actividad principal.
- **Principal:** Evidencia más apta a partir de los 3 análisis.
- **Valor:** Valor numérico que mide el grado de relación entre ambas *frases*.
- **Apto:** Escala de 0 a 100, donde 100 indica una relación apta.
- **Secundario:** Información extra de actividad alternativa encontrada.

4.2.6- Relación entre Evidencias e Instrumentos

Las evidencias que el alumno entrega al cumplir con una actividad, necesitan de una forma de evaluación que se adecúe a ellas. Como se muestra en la **Figura 29**, se realiza un recorrido por las evidencias teniendo el objeto didáctico identificado y se asocia si es apto o no el instrumento seleccionado en base a la información almacenada en la ontología.

```
Por pos, frase en Enumeración (Lista Evidencias):  
instrumentos = Consultar Ontología Instrumentos Evidencia(frase['actividad'])  
instrumentos = Consultar Ontología Sinónimo Instrumentos(instrumentos)  
  
Por ins en Instrumentos:  
valor = Distancia Levenshtein (ins, lista Instrumentos[pos])  
Si valor es mayor a 90:  
    encontrado = True  
    break
```

Figura 29-Seudocodigo base de evidencias e instrumentos

Como resultado de esta fase del análisis, se retorna un JSON que contiene los siguientes parámetros:

- **Encontrado:** Variable booleana que indica si el match ha sido exitoso.
- **Evidencia:** Actividad que se busca relacionar.
- **Instrumento:** Instrumento que será utilizado para realizar la evaluación.

Capítulo 5-Análisis de resultados

En el presente capítulo se analizan los resultados obtenidos a partir de la implementación de la ontología y el algoritmo de PLN. El capítulo se segmenta en tres partes, en la primera se analiza una de las Instrumentaciones Didácticas (ID) escritas por docentes de educación superior que no han tomado el curso previo de la plataforma. La ventaja que representa el analizar este tipo de ID es que ayuda a detectar errores habituales en la redacción y excepciones al patrón, situaciones que se esperan erradicar con el curso de capacitación y formación docente. También, se considera un análisis del JSON resultante por cada proceso realizado, el cual es enviado al módulo de recomendación para que realice las recomendaciones pertinentes.

Posteriormente, se analiza a detalle los resultados obtenidos de una de las ID escritas utilizando la metodología planteada en la presente investigación, siendo una de las pruebas piloto que permiten identificar si se tiene o no una mejora en los resultados basándose en la comparativa en el cumplimiento del patrón siguiendo la metodología del modelo EBC.

Por último, se presentan estadísticas del desempeño del algoritmo en comparativa con los resultados de uno de los evaluadores del departamento encargado de la labor, lo cual da lugar a una discusión de los resultados.

5.1.-ID-Prueba I

Los datos importantes de la unidad que necesita el módulo para realizar el análisis son: la competencia específica, el temario, las actividades de aprendizaje, las actividades de enseñanza, las evidencias de aprendizaje y los instrumentos de evaluación. Para esta prueba, se utilizó una ID de la carrera de Ingeniería en Informática para el ciclo escolar 2018-2019. Los datos de la

Tabla 7 corresponden a la primera unidad de dicha ID.

Tabla 7-Instrumentación didáctica de prueba

Competencia: Identifica los fundamentos de la interacción usuario-computadora, a partir del conocimiento del proceso cognitivo humano y los principios de diseño de objetos.		
Temario	Actividades de aprendizaje	Actividades de enseñanza
-Proceso Cognitivo Humano -Proceso de Aprendizaje Humano -Ergonomía cognitiva -Conceptos básicos -Fundamentos del diseño -Principios de diseño de objetos.	-Mapa conceptual de términos básicos de la interacción usuario-computadora. -Análisis de estudios de caso de interfaces de usuario. -Ejercicios de aplicación de los principios de diseño al diseño de objetos virtuales. -Elaboración de documento con definición del proyecto de interfaz de usuario.	-Introducción a los conceptos y principios del diseño, mediante la elaboración de una presentación en Power Point y el uso de ejemplos. -Presentación de estudios de caso. -Formulación de preguntas.
Evidencias de aprendizaje		Instrumentos de evaluación
-Instrumento de recolección de datos -Reporte de análisis de datos y resultados -Documento con establecimiento de requerimientos		-Lista de cotejo -Rúbrica -Rúbrica

5.2.- Pruebas al autómata

Como se mencionó anteriormente, es necesario segmentar la competencia haciendo uso del autómata diseñado. Para el texto “*Identifica los fundamentos de la interacción usuario-computadora, a partir del conocimiento del proceso cognitivo humano y los principios de diseño de objetos*”, el algoritmo da como resultado la segmentación de la **Figura 30**. Como se aprecia, la competencia contiene los tres elementos necesarios para ser considerada válida, por lo que no es enviada a revisión. De igual forma, es posible determinar

el tipo de elemento que contiene en su predicado, siendo identificado como una condición de ejecución. Para los casos en los cuales no es posible determinar su tipo, el valor retornado es *NA* y eso merita una revisión por parte del evaluador.

```
Original:Identifica los fundamentos de la interacción usuario-computadora, a partir del
conocimiento del proceso cognitivo humano y los principios de diseño de objetos.
Revisar-Opcional:False-False
Verbo:identificar
Objeto:el fundamento de el interacción usuario_computador ,
Elemento 3 de tipo C (-): a_partir_de el conocimiento de el proceso cognitivo humano y
el principio de diseño de objeto .
```

Figura 30-Resultado análisis de autómeta

Tomando al verbo *identificar*, se realizó la consulta a la ontología para obtener su dominio y actividades asociadas dando como resultado un agrupamiento en los niveles de *análisis*, *comprensión* y *conocimiento*. Como se observa en la **Figura 31**, se recuperaron 152 actividades de aprendizaje y evidencias asociadas al dominio, con 169 verbos posibles a utilizar, 61 actividades de enseñanza y 48 instrumentos como predeterminados. Cabe resaltar que, antes de la recuperación se realizó un razonamiento deductivo para inferir conocimiento a partir de las relaciones ya existentes en el modelo.

```
Dominio: Analisis, Comprension, Conocimiento
Verbos:169
Actividades de aprendizaje: 152
Actividades de Enseñanza: 61
Evidencias de aprendizaje:152
Instrumentos de evaluacion:48
```

Figura 31-Actividades asociadas en base al dominio de la competencia

5.3.-Pruebas Match

Como siguiente paso, se realizó el proceso de match entre las actividades de aprendizaje y el tercer elemento de la competencia devolviendo como resultado las variables mostradas en la **Figura 32**, lo que indica que dicha actividad es la que permitirá cumplir con la competencia planteada para esa unidad. Cabe destacar que este paso considera valido el match si se encuentra al menos una actividad relacionada, debido a que pueden llevar una secuencia entre ellas dando una infinidad de posibles combinaciones o su relación no esté explícitamente escrita, por lo que el resto de la comparativa lo analiza el evaluador final.

```
Encontrado: True
mapa conceptual término básico interacción usuario computador
ejercicio aplicación principio diseño diseño objeto virtual
```

Figura 32-Match entre actividad de aprendizaje y competencia

Posteriormente, se busca el match entre las actividades recuperadas de la ontología y las redactadas, considerando si se ha utilizado el verbo de dominio para acompañarle y el tema que se desea abordar. La **Tabla 8** muestra el resultado del match de las actividades de aprendizaje. Con el fin de agilizar el proceso, se realiza la búsqueda de la actividad, después el verbo asociado a ella y por último el tema. Si alguno de los elementos no es identificado, se retorna *NA* en vez de la cadena vacía.

Tabla 8-Match de actividades de aprendizaje

Temario	Actividad de aprendizaje	Algoritmo
-Proceso Cognitivo Humano -Proceso de Aprendizaje Humano -Ergonomía cognitiva -Conceptos básicos -Fundamentos del diseño -Principios de diseño de objetos.	Mapa conceptual de términos básicos de la interacción usuario-computadora.	Revisar Actividad: <i>NA</i> mapa conceptual Tema: <i>NA</i>
	Análisis de estudios de caso de interfaces de usuario.	Revisar Actividad: <i>Análisis</i> estudio de caso Tema: <i>NA</i>
	Ejercicios de aplicación de los principios de diseño al diseño de objetos virtuales.	Revisar Actividad: <i>NA</i> Ejercicios Tema: Principio de diseño de objetos
	Elaboración de documento con definición del proyecto de interfaz de usuario	Revisar Actividad: <i>Elaboración</i> documento Tema: <i>NA</i> <hr/> Actividad: <i>Definición</i> proyecto

El proceso de análisis realizado se repite para las actividades de enseñanza, evidencias de aprendizaje e instrumentos de evaluación, cuyo resultado es mostrado en la **Tabla 9**, **Tabla 10** y **Tabla 11** respectivamente. Puntos importantes por mencionar son que en la **Tabla 9** se tienen asociados dos temas a una misma actividad, mientras que el resto de ellas lo tienen aun cuando cumple con los dos primeros elementos en la redacción.

Tabla 9-Match de actividades de enseñanza

Actividades de Enseñanza	Algoritmo
Introducción a los conceptos y principios del diseño, mediante la elaboración de una presentación en Power Point y el uso de ejemplos.	OK Actividad: <i>Elaboración</i> presentación Tema: Principios de diseño de objetos
Presentación de estudios de caso.	Revisar Actividad: <i>Presentación</i> estudio de caso Tema: NA
Formulación de preguntas.	Revisar Actividad: <i>Formulación</i> pregunta Tema: NA

En cuanto a la **Tabla 10** se mantiene el error en la primera actividad, aunque no es necesario acompañarla de un verbo y en consecuencia no se somete a revisión. La única actividad que si es revisada es la segunda al carecer de un tema asociado, el cual fue redactado en las actividades de aprendizaje, pero este no fue considerado tanto en las actividades de enseñanza como en las evidencias, lo que puede determinar un patrón en la secuencia.

Tabla 10-Match de evidencias de aprendizaje

Evidencias de aprendizaje	Algoritmo
Mapa conceptual de términos abordados.	OK Actividad: <i>NA</i> mapa conceptual Tema: Conceptos básicos
Reporte de estudio de caso	Revisar Actividad: <i>Reporte</i> estudio de caso Tema: NA
Ejercicios de diseño de objetos virtuales.	OK Actividad: <i>NA</i> ejercicio Tema: Principios de diseño de objetos

Finalmente, la **Tabla 11** indica únicamente el instrumento que se ha encontrado al carecer (a diferencia de los anteriores) de un verbo y un tema que le acompañe.

Tabla 11-Match de instrumentos de evaluación

Instrumentos de evaluación	Algoritmo
Rúbrica	OK
Lista de control	OK
Lista de control	OK

5.4.-Prueba de relación entre actividades y evidencias

Con el material didáctico identificado, es necesario encontrar si existe relación entre actividades y evidencias. En algunos casos, se necesita de un análisis bidireccional debido a que la cantidad de elementos en ambos grupos no es la misma, tal como el resultado mostrado en la **Tabla 12** que contiene cuatro actividades y solo tres posibles evidencias, por lo que una de ellas no tiene un elemento asociado. Caso contrario, la **Tabla 13** indica que todas las evidencias se relacionan con una actividad.

Tabla 12-Evidencias a partir de actividades

Actividad	Evidencia	Apto
Mapa conceptual de términos básicos de la interacción usuario computadora.	Mapa conceptual de términos abordados.	OK
Análisis de estudios de caso de interfaces de usuario.	Reporte de estudio de caso.	OK
Ejercicios de aplicación de los principios de diseño al diseño de objetos virtuales.	Ejercicios de diseño de objetos virtuales.	OK
Elaboración de documento con definición del proyecto de interfaz de usuario.	NA	Revisar

Tabla 13-Actividades a partir de evidencias

Evidencia	Actividad	Apto
Mapa conceptual de términos. abordados.	Elaborar mapa conceptual de términos básicos de la interacción usuario-computadora.	OK
Reporte de estudio de caso.	Análisis de estudios de caso de interfaces de usuario.	OK
Ejercicios de diseño de objetos virtuales.	Ejercicios de aplicación de los principios de diseño al diseño de objetos virtuales.	OK

5.5.-Prueba de evaluación de evidencias

Finalmente, se procedió a buscar si el instrumento elegido para evaluar las evidencias redactadas es el adecuado. En base al ejemplo mostrado en la **Tabla 14**, la actividad *ejercicio* se indica que es incorrecto evaluarla con una *lista de control* por lo que se retroalimenta con esta indicación al módulo de recomendación. En este punto termina el proceso, el cual es reiterativo para cada unidad.

Tabla 14-Instrumento de evaluación de evidencias

Evidencias	Instrumentos	Algoritmo
Mapa conceptual de términos abordados.	Rúbrica	OK Actividad: Mapa conceptual
Reporte de estudio de caso	Lista de control	OK Actividad: Estudio de caso
Ejercicios de diseño de objetos virtuales.	Lista de control	Revisar Actividad: Ejercicio

5.6.-Resultados enviados al módulo de recomendación

Las siguientes tablas muestran algunas de las variables resultantes que son enviadas al módulo de recomendación, así como una breve descripción sobre que les hace relevantes en el análisis. La **Tabla 15** contiene las variables de la competencia específica, donde se subdividen los componentes enfocándose en sus tres elementos básicos. Para notificar en caso de error, es enviado un valor booleano que avisa si es necesario realizar una revisión o ha cumplido satisfactoriamente con la metodología.

Tabla 15-Variables resultantes de competencia específica

Nombre	Valor	Descripción
Revisar	True	Notifica si la revisión ha sido exitosa.
Ontología	True	Notifica si pertenece a la ontología.
Verbo	Identifica	Ayuda a determinar el dominio.
Objeto	El fundamento de la interacción usuario computadora	Objeto de la competencia.
Tipo	Condición	Identificación del tercer elemento como condición de ejecución o finalidad.
Elemento	A partir del conocimiento del proceso cognitivo humano y el principio de diseño de objeto	Tercer elemento de la competencia.
Corregir	False	Indica si se necesita corregir el verbo.

Mientras tanto, la **Tabla 16** contiene el resultado del match entre la competencia y las actividades de aprendizaje. En caso de tener más de una relacionada, se van agregando vectores siguiendo el mismo esquema de la fila de *actividad*.

Tabla 16-Variables resultantes entre competencia y actividades

Nombre	Valor	Descripción
Encontrado	True	Indica si existe relación.
Competencia	El fundamento de la interacción usuario computadora, a partir del...	Considera solo el objeto y condición de ejecución.
Actividad	Mapa conceptual de términos básicos de la interacción usuario computadora.	<i>Frase</i> asociada.
Similaridad	67	Valor de similitud.

Posteriormente, la **Tabla 17** retorna los valores del match interno, centrando la estructura para cada una de ellas en los tres elementos principales, considerando un valor booleano fácil de identificar que indica si ha sido encontrado y si pertenece al dominio.

Tabla 17-Variables resultantes del análisis del Match

Nombre	Valor	Descripción
Original	Mapa conceptual términos básicos interacción usuario-computadora	Actividad original analizada.
Lematizada	Mapa conceptual término básico interacción usuario computador	Texto pre procesado y con cada token en su forma base.
Actividad encontrada	True	Booleano que indica si se ha encontrado la actividad.
Actividad	Mapa conceptual	Actividad encontrada.
Verbo encontrado	False	Booleano que indica si se ha encontrado el verbo.
Verbo	NA	Verbo que acompaña a la actividad.
Tema encontrado	False	Booleano que indica si se ha encontrado un tema asociado a la <i>frase</i> .
Tema	NA	Tema asociado.

En cuanto a la **Tabla 18**, se indican los valores resultantes para la relación entre las actividades de aprendizaje y sus evidencias. A pesar del resultado final depende de 3 diferentes métodos, la evaluación de si es apta la evidencia dada una actividad o viceversa se retorna con un valor numérico en escala del 0 al 100 donde 100 indica una alta correlación. Por otra parte, en valores menores a 60 notifica que es necesaria una revisión y uno inferior a 30 es considerado una relación inexistente. La estructura es repetitiva para cada *frase*.

Tabla 18-Variables resultantes de las actividades y evidencias

Nombre	Valor	Descripción
Actividad	Mapa conceptual término básico interacción usuario computador	Actividad base.
Principal	Mapa conceptual termino abordar	Evidencia principal relacionada.
Unión	Mapa conceptual	Actividad existente en ambas <i>frases</i> .
Secundario	NA	Posible segunda evidencia.
Distancia	84	Valor escalar de similaridad.
Apto	100	Indica que tan apta es la relación.

Finalmente, el diccionario retornado para informar si el instrumento es apto para evaluar la evidencia está basado en el esquema de la **Tabla 19**, siendo el ultimo obtenido para una única unidad de la ID.

Tabla 19-Variables resultantes entre instrumentos y evidencias

Nombre	Valor	Función
Apto	True	Booleano que indica si es apta su correlación.
Evidencia	Mapa conceptual	Evidencia que se pretende evaluar.
Instrumento	Rúbrica	Instrumento por utilizar para la evaluación.

El motivo por el cual están segmentados en diferentes procesos es para dar la oportunidad al docente de ir evaluando su ID conforme va avanzando y no sea necesario tenerla terminada para su análisis. No obstante, la secuencia que se sigue en el ciclo es muy importante al depender cada proceso del anterior, partiendo desde la competencia específica hasta la evaluación de las evidencias de la unidad.

Ahora bien, analizando el resto de las unidades que componen esta ID, se encontró la existencia de un patrón en los errores de redacción cometidos, recayendo principalmente en la omisión de un verbo que acompañe al objeto didáctico que ha seleccionado. De igual forma, algunas de las *frases* analizadas carecen de un subtema de la unidad o hacen relación a uno de alguna unidad diferente, por lo que son enviadas a revisión al no cumplir con el patrón esperado. Finalmente, aunque no existe relación entre algunas actividades de aprendizaje y evidencias de aprendizaje, la mayor parte de los errores recaen en la asociación de las actividades de aprendizaje con la competencia, debido a que trata de unirlas en secuencia o implícitamente entre ellas en vez de crear un enlace directo con la competencia,

por lo que la mayor parte de los elementos de este tipo son enviados a revisión tanto por el algoritmo como por el evaluador.

5.7.-ID-Prueba piloto

La prueba anterior permitió el confirmar algunas excepciones a la regla cuando no se hace la redacción de acuerdo con la metodología propuesta en el Modelo EBC, por lo que la cantidad de elementos enviados a revisión es alta. Para obtener una comparativa entre los resultados, se optó por una prueba piloto en la cual se analizaron ID que cumplieran con la metodología de la cual se obtuvieron los siguientes datos.

La **Tabla 20** contiene una de las unidades de la ID resultante de la prueba piloto realizada. En ella, se puede observar una mayor cantidad de *frases* que cumplen con el patrón de un verbo acompañando a la actividad y uno de los temas del temario que les identifique. Respecto a los resultados del algoritmo, se analizó la competencia específica de la unidad obteniendo la siguiente segmentación:

- **Verbo:** Identificar
- **Objeto:** Las diferentes estructuras de datos, su clasificación y forma de manipularlas
- **Finalidad:** para buscar la manera más eficiente de resolver problemas
- **Condición:** A través de su clasificación y manejo de memoria.

En definitiva, el autómata identificó las cuatro secciones que componen la ID lo cual es lo óptimo en la redacción ya que, aunque la finalidad puede ser considerada como opcional, la existencia de ambos elementos ayuda a relacionar el por qué y para que cumplir la competencia específica de la unidad, lo que otorga más información para relacionarlo con las actividades de aprendizaje, correspondencia que se observa en la **Tabla 21**, donde el algoritmo identificó una alta relación en solo una de las actividades, mientras que 4 de ellas se obtuvo una dependencia parcial suficiente para ser considerada apta. Por otra parte, el algoritmo y el evaluador difirieron en la actividad “*Elaborar cuadro comparativo del manejo de memoria estática y dinámica*”, donde en la competencia la parte de “*estática y dinámica*” está implícito por lo que el algoritmo no logra identificarlo.

Tabla 20-Instrumentación didáctica de prueba piloto

Competencia: Identifica las diferentes estructuras de datos, su clasificación y forma de manipularlas para buscar la manera más eficiente de resolver problemas a través de su clasificación y manejo de memoria.		
Temario	Actividades de aprendizaje	Actividades de enseñanza
Introducción a las estructuras de datos 1.1 Clasificación de las estructuras de datos 1.2 Tipos de datos abstractos (TDA) 1.3 Ejemplos de TDA's 1.4 Manejo de memoria 1.4.1 Memoria estática 1.4.2 Memoria dinámica 1.5 Análisis de algoritmos 1.5.1 Complejidad en el tiempo 1.5.2 Complejidad en el espacio 1.5.3 Eficiencia de los algoritmos	Elaborar cuadro sinóptico de las diferentes estructuras de datos y su clasificación. Elaborar resumen sobre tipos de datos abstractos e implementación en la programación orientada a objetos. Elaborar cuadro comparativo del manejo de memoria estática y dinámica. Realizar práctica de laboratorio para medir la eficiencia de algoritmos en tiempo de ejecución. Realizar práctica de laboratorio para el uso de la memoria dinámica. Elaborar cuadro comparativo de dos algoritmos que resuelvan el mismo problema.	Uso de TIC para la búsqueda de información acerca de las estructuras de datos. Organizar debate entre los estudiantes acerca de los tipos de memoria. Presentar problemas acerca del uso de vectores en la programación. Presentar problemas acerca de eficiencia en la programación
Evidencias de aprendizaje		Instrumentos de evaluación
Resumen de tipos de datos abstractos Prácticas de memoria estática y memoria dinámica Cuadro comparativo de eficiencia de algoritmos de estructuras de datos para un mismo problema		Portafolio de evidencias Portafolio de evidencias Portafolio de evidencias

Consecutivamente, se obtuvieron los datos del match de las actividades de aprendizaje (**Tabla 22**), actividades de enseñanza (**Tabla 23**), evidencias de aprendizaje (**Tabla 24**) e instrumentos de evaluación (**Tabla 25**). Para la **Tabla 22** las seis frases cumplen con el patrón establecido al contener un verbo que acompaña a cada una de las actividades propuestas que a su vez están relacionadas con al menos uno de los subtemas de la unidad.

Tabla 21-Match entre competencias y actividades (Prueba piloto)

Nombre	Valor
Encontrado	True
Competencia	Identifica las diferentes estructuras de datos, su clasificación y forma de manipularlas para buscar la manera más eficiente de resolver problemas a través de su clasificación y manejo de memoria.
Actividad principal	<ul style="list-style-type: none"> -Elaborar cuadro sinóptico de las diferentes estructuras de datos y su clasificación. -Elaborar resumen sobre tipos de datos abstractos e implementación en la programación orientada a objetos. -Elaborar cuadro comparativo del manejo de memoria estática y dinámica. -Realizar práctica de laboratorio para medir la eficiencia de algoritmos en tiempo de ejecución. -Realizar práctica de laboratorio para el uso de la memoria dinámica. -Elaborar cuadro comparativo de dos algoritmos que resuelvan el mismo problema.
Similaridad máxima	75

Otro punto para considerar es que el algoritmo ha identificado que el docente utilizó al menos una actividad para cada subtema, eso sin considerar aquellos subtemas que a su vez se subdividen. Si se realiza una comparativa en la redacción de la ID sin la metodología con respecto a esta, el cambio es notorio al poderse identificar de forma clara y precisa a los elementos y su correlación con el resto de la ID.

Tabla 22-Match actividades de aprendizaje (Prueba piloto)

Actividad de aprendizaje	Algoritmo
Elaborar cuadro sinóptico de las diferentes estructuras de datos y su clasificación.	OK Actividad: <i>Elaborar</i> cuadro sinóptico Tema: Clasificación estructura de datos
Elaborar resumen sobre tipos de datos abstractos e implementación en la programación orientada a objetos.	OK Actividad: <i>Elaborar</i> resumen Tema: Tipo de dato abstracto
Elaborar cuadro comparativo del manejo de memoria estática y dinámica.	OK Actividad: <i>Elaborar</i> cuadro comparativo Tema: Manejo de memoria
Realizar práctica de laboratorio para medir la eficiencia de algoritmos en tiempo de ejecución.	OK Actividad: <i>Realizar</i> practica Tema: Eficiencia de algoritmo
Realizar práctica de laboratorio para el uso de la memoria dinámica.	OK Actividad: <i>Realizar</i> práctica de laboratorio Tema: Manejo de memoria
Elaborar cuadro comparativo de dos algoritmos que resuelvan el mismo problema.	OK Actividad: <i>Elaborar</i> cuadro comparativo Tema: Análisis de algoritmos

Continuando con el análisis, la **Tabla 23** contiene a la única *frase* que es enviada a revisión debido a la carencia de un tema de la asignatura que le acompañe ante la falta de conocimiento tácito que permita identificar a los vectores como estructura de datos, las cuales son el tema central de la unidad. Por otra parte, el resto de las actividades cumplen con el patrón esperado.

Tabla 23-Match de actividades de enseñanza (Prueba piloto)

Actividades de Enseñanza	Algoritmo
Uso de TIC para la búsqueda de información acerca de las estructuras de datos.	OK Actividad: <i>Búsqueda</i> de información Tema: Introducción a la estructura de datos
Organizar debate entre los estudiantes acerca de los tipos de memoria.	OK Actividad: <i>Organizar</i> debate Tema: Manejo de memoria
Presentar problemas acerca del uso de vectores en la programación.	Revisar Actividad: <i>Presentar</i> problema Tema: NA
Presentar problemas acerca de eficiencia en la programación	OK Actividad: <i>Presentar</i> problema Tema: Eficiencia algoritmo

Acto seguido, el resultado de la **Tabla 24** contiene la excepción de que el verbo puede ser opcional en la redacción de las evidencias de aprendizaje, por lo que cumple con el patrón establecido para esta sección. Cabe resaltar que la evidencia de *cuadro comparativo de eficiencia de algoritmos de estructuras de datos para un mismo problema* está relacionada con 4 de los subtemas de la unidad.

Tabla 24-Match de evidencias de aprendizaje (Prueba piloto)

Evidencias de aprendizaje	Algoritmo
Resumen de tipos de datos abstractos	OK Actividad: NA resumen Tema: Tipos de datos abstractos
Prácticas de memoria estática y memoria dinámica	OK Actividad: NA practicas Tema: Memoria dinámica
Cuadro comparativo de eficiencia de algoritmos de estructuras de datos para un mismo problema	OK Actividad: NA cuadro comparativo Tema: Eficiencia de algoritmo

Finalmente, la **Tabla 25** identifica a los instrumentos que han sido elegidos para la evaluación de las evidencias, donde llama la atención que se ha seleccionado el mismo instrumento para evaluar distintas evidencias de aprendizaje, lo cual puede ser considerado por el módulo de recomendación para otorgar una retroalimentación.

Tabla 25-Match de instrumentos de evaluación (Prueba piloto)

Instrumentos de evaluación	Algoritmo
Portafolio de evidencias	OK
Portafolio de evidencias	OK
Portafolio de evidencias	OK

Por otra parte, la **Tabla 26** contiene la relación entre las actividades de aprendizaje y las evidencias donde, a pesar de ser una comparativa entre 6 actividades para solo tres evidencias, se encontró una correspondencia en cada una de ellas. Un dato por considerar es la primera relación donde haciendo uso de los sinonimos fue posible identificar que un *cuadro sinóptico* puede ser un *cuadro comparativo*, este conocimiento inferido es debido a que ambos elementos están relacionados en el modelo de conocimiento por la palabra *diagrama*.

Tabla 26-Evidencias asociadas a actividades de aprendizaje (Prueba piloto)

Actividad	Evidencia	Apto
Elaborar cuadro sinóptico de las diferentes estructuras de datos y su clasificación.	Cuadro comparativo de eficiencia de algoritmos...	OK
Elaborar resumen sobre tipos de datos abstractos e implementación en la programación orientada a objetos.	Resumen de tipos de datos abstractos	OK#
Elaborar cuadro comparativo del manejo de memoria estática y dinámica.	Prácticas de memoria estática y memoria dinámica	OK#
Realizar práctica de laboratorio para medir la eficiencia de algoritmos en tiempo de ejecución.	Cuadro comparativo de eficiencia de algoritmos...	OK
Realizar práctica de laboratorio para el uso de la memoria dinámica.	Prácticas de memoria estática y memoria dinámica	OK
Elaborar cuadro comparativo de dos algoritmos que resuelvan el mismo problema.	Cuadro comparativo de eficiencia de algoritmos de estructuras de datos para un mismo problema	OK#

Por otra parte, este es un claro ejemplo de la factibilidad de hacer uso de la Distancia de Levenshtein para encontrar la similaridad entre las actividades y las evidencias al poder identificarse los elementos como si fuesen una paráfrasis de su contraparte, donde su valor de similitud es alto si se compara con el obtenido de la relación con el resto de las frases del otro grupo.

Por último, la **Tabla 27** indica que a pesar de que solamente se ha utilizado un instrumento para calificar las evidencias, este es óptimo para hacer su evaluación.

Tabla 27- Instrumentos asociados a evidencias de aprendizaje (Prueba piloto)

Evidencias	Instrumentos	Algoritmo
Resumen de tipos de datos abstractos	Portafolio de evidencias	OK Actividad: Resumen
Prácticas de memoria estática y memoria dinámica	Portafolio de evidencias	OK Actividad: Práctica
Cuadro comparativo de eficiencia de algoritmos de estructuras de datos para un mismo problema	Portafolio de evidencias	OK Actividad: Cuadro comparativo

Como conclusión de la prueba piloto, se identificó un cambio en el patrón de redacción de ID sin seguir la metodología del modelo EBC y aquellas que si lo siguieron, lo cual tuvo como consecuencia una disminución de la cantidad de *frases* que eran enviadas a revisión. Uno de los cambios más significativos fue que el docente utilizó adecuadamente los verbos para acompañar a las actividades y estas las relacionó con los temas de la unidad actual sin intentar abarcar los de unidades anteriores o posteriores. De igual forma, la relación entre los elementos de las distintas secciones podía identificarse mejor, lo cual se vio reflejado en un aumento considerable en las *frases* que eran consideradas correctas y una disminución de las enviadas a revisión.

5.8-Comparativa de resultados

Para evaluar el desempeño del algoritmo, se realizó una comparativa entre sus resultados y los de uno de los expertos que aportó su conocimiento al modelo ontológico. Primeramente, la **Tabla 28** muestra que ambos coinciden en que la competencia esta correctamente redactada al incluir un verbo que permita identificar su nivel cognitivo en base a la taxonomía de Bloom.

Tabla 28- Comparativa de competencia específica

Competencia	Evalgador	Algoritmo
Diseña interfaces de usuario gráficas, hápticas y gestuales para dispositivos móviles, identificando sus características principales.	Diseña interfaces de usuario gráficas, hápticas y gestuales para dispositivos móviles, identificando sus características principales.	OK Verbo: Diseña Objeto: interfaces de usuario gráficas, hápticas y gestuales Finalidad: para dispositivos móviles, Condición: identificando sus características principales. Corregir verbo: Si Encontrado en la Ontología: Si Escrito en Infinitivo: No
Nota del evaluador: OK		

Ahora bien, los datos suplementarios obtenidos del autómata no hacen posible obtener una relación con las actividades de aprendizaje y el evaluador, motivo por el que ambos indican la inexistencia de una dependencia directa con la competencia. La **Tabla 29** muestra

que el evaluador y el algoritmo identificaron a dos de las cuatro actividades con una codependencia, sin embargo, en una de ellas se difiere debido al manejo del conocimiento empírico donde el evaluador considera que Android e iOS son temas asociados a dispositivos móviles, por lo que esta actividad no es enviada a revisión.

Tabla 29- Comparativa de relación entre competencia y actividades

Actividades de Aprendizaje	Evaluador	Algoritmo
Elaboración de estudio comparativo de la interfaz de usuario de tres apps.	OK	OK
Elaboración del Viaje de usuario (User Journey).	Revisar	Revisar
Elaboración de tabla comparativa de características de interfaces de usuario para Android e iOS.	OK	OK
Diseño de bocetos (Wireframe) para Android e iOS.	OK	Revisar
Nota del evaluador: Revisar		

Posterior a eso, la **Tabla 30** muestra que tanto el algoritmo como el evaluador han coincidido que la primera actividad no necesita una revisión y las dos siguientes les hace falta un tema de la unidad. A pesar de que *viaje de usuario* es un tema de la asignatura, este corresponde a una unidad diferente, motivo por el cual no es considerado como válido para el presente ejemplo.

Tabla 30- Comparativa en match de actividades de enseñanza

Actividades de Enseñanza	Evaluador	Algoritmo
Realizar exposición de introducción, en Power Point, a los principios básicos de usabilidad en el diseño de interfaces de usuario.	OK	OK Actividad: Realizar exposición Tema: Los principios de usabilidad aplicado al diseño de interfaces de usuario.
Demostración, con ejemplos, de cómo se elabora un viaje de usuario.	Falta tema.	Revisar Actividad: Demostración ejemplo Tema: NA
Demostración, con ejemplos, de cómo se diseña un boceto de app.	Falta tema.	Revisar Actividad: Demostración ejemplo Tema: NA
Nota del evaluador: Las actividades no tienen un tema relacionado.		

Acto seguido, la **Tabla 31** contiene una problemática idéntica a la anterior, donde la falta de un tema asociado para la segunda actividad sea considerada para una revisión, coincidiendo el evaluador y el algoritmo en esta decisión. Por otra parte, en la última

actividad difieren el evaluador y el algoritmo debido a que Android e iOS son sistemas operativos de dispositivos móviles, tema existente en la unidad y que por segunda ocasión implica una actividad enviada a revisión.

Tabla 31- Comparativa en match de actividades de aprendizaje

Actividad de aprendizaje	Evaluador	Algoritmo
Elaboración de estudio comparativo de la interfaz de usuario de tres apps.	OK	OK Actividad: <i>Elaboración</i> estudio comparativo Tema: La affordance aplicada al diseño de interfaces de usuario
Elaboración del viaje de usuario (User Journey).	Falta tema.	Revisar Actividad: <i>Elaboración</i> viaje de usuario Tema: NA
Elaboración de tabla comparativa características de interfaces de usuario para Android e iOS.	OK	OK Actividad: <i>Elaboración</i> tabla comparativa Tema: La affordance aplicada al diseño de interfaces de usuario
Diseño de bocetos (Wireframe) para Android e iOS.	OK	Revisar Actividad: <i>Diseño</i> Bocetos Tema: NA
Nota del evaluador: Falta relacionar actividad con los temas de la unidad.		

Continuando con el proceso de match, la **Tabla 32** muestra deficiencias en todos sus elementos al solo especificar la actividad y no un tema asociado en todos sus elementos. Cabe resaltar que el problema sobre el conocimiento empírico se mantiene y eso queda demostrado en la única actividad en la cual se mantiene la discrepancia entre el evaluador y el algoritmo lo cual se puede considerar como un patrón.

Tabla 32- Comparativa en match de evidencias de aprendizaje

Evidencias de aprendizaje	Evaluador	Algoritmo
Estudio comparativo de tres apps.	Falta tema.	Revisar Actividad: NA Estudio comparativo
Viaje de usuario	Falta tema.	Revisar Actividad: NA viaje de usuario
Tabla comparativa de Android e iOS	OK	Revisar Actividad: NA tabla comparativa Tema: NA
Boceto de app	Falta tema.	Revisar Actividad: NA boceto
Nota del evaluador: Falta redactar el tema asociado.		

Para finalizar este paso, se busca que los instrumentos seleccionados coincidieran con los predeterminados obteniendo el resultado de la **Tabla 33**, donde tanto el evaluador como el algoritmo concuerdan en que cada uno de los elementos son correctos.

Tabla 33- Comparativa en match de instrumentos de evaluación

Instrumentos de evaluación	Evaluador	Algoritmo
Rúbrica	OK	OK
Lista de verificación	OK	OK
Rúbrica	OK	OK
Rúbrica	OK	OK
Nota del evaluador: OK.		

Por otra parte, para cada actividad es necesaria una evidencia y esto se demuestra en la **Tabla 34** donde la comparativa indica una relación entre los elementos siendo considerados como aptos. En esta ocasión, tanto el evaluador como el algoritmo coinciden en que todos los elementos están relacionados entre sí al existir un match del objeto de aprendizaje. A pesar de ello, la redacción de ambos grupos es deficiente, por lo que es necesario hacer las modificaciones correspondientes y volver a evaluar.

Tabla 34- Comparativa de relación entre actividades y evidencias

Actividades de Aprendizaje	Evaluador	Algoritmo
Elaboración de estudio comparativo de la interfaz de usuario de tres apps.	OK Estudio comparativo de tres apps.	OK Actividad: Estudio comparativo de tres apps. Match: Estudio comparativo
Elaboración del viaje de usuario (User Journey).	OK Viaje de usuario	OK Actividad: Viaje de usuario Similar: Viaje de usuario
Elaboración de tabla comparativa de características de interfaces de usuario para Android e iOS.	OK Tabla comparativa de Android e iOS	OK Actividad: Tabla comparativa de Android e iOS Similar: Tabla comparativa
Diseño de bocetos (Wireframe) para Android e iOS.	OK Boceto de app	OK Actividad: Boceto de app Similar: Boceto
Nota del evaluador: OK		

Finalmente, la **Tabla 35** muestra que todos los instrumentos que se han elegido para evaluar a las evidencias son aptos desde el punto de vista del evaluador y del método implementado, correlación que se obtiene del conocimiento almacenado en la ontología y

que permite hacer uso de la inferencia para obtener los instrumentos adecuados para la evaluación de las evidencias encontradas durante el proceso de match.

Tabla 35- Comparativa de relación entre evidencias e instrumentos

Evidencias	Instrumentos	Evalua dor	Algoritmo
Estudio comparativo de tres apps.	Rúbrica	OK	OK Actividad: Estudio comparativo
Viaje de usuario	Lista de verificación	OK	OK Actividad: Viaje de usuario
Tabla comparativa de Android e iOS	Rúbrica	OK	OK Actividad: Tabla comparativa
Boceto de app	Rúbrica	OK	OK Actividad: Boceto
Nota del evaluador: OK			

5.9-Discusión de resultados

Basándose en los resultados mostrados, se encontró un patrón a seguir en los errores cometidos por los docentes quienes no acompañan a los objetos de un verbo que permita determinar el nivel cognitivo que se pretende alcanzar al cumplir con dicha actividad. A la vez, algunas *frases* redactadas carecen de un tema de la unidad que fundamenten su utilización aun cuando los elementos están relacionados entre sí, siendo un error común el utilizar subtemas asociados a una unidad anterior o posterior a la cual se está redactando.

La Tabla 36 contiene los datos de los 715 elementos analizados durante las pruebas donde se coincidió en 658 (92.03%) de ellos de los cuales 496 (75.37%) se consideró que cumplían con el patrón de redacción y 162 (24.63%) de ellos fueron enviados directamente a revisión debido a la inexistencia de un verbo o subtema para acompañar al objeto didáctico. Como punto negativo, en 57 (7.97%) de los elementos analizados, el algoritmo y el evaluador obtuvieron resultados diferentes de los cuales 29 (50.87%) se dieron en la relación entre la competencia específica y las actividades de aprendizaje. Cabe resaltar que la mayor parte de estas divergencias se dieron al incluir el conocimiento empírico del evaluador debido a que en el texto la correlación no estaba explícitamente escrita y era necesario inferir conocimiento muy específico que no fue incluido en la ontología. Además, 6 (46.15%) de las 13 competencias que se analizaron fueron enviadas a revisión tanto por el algoritmo como el

evaluador lo cual indica que un patrón de errores que se van formando desde la redacción de la competencia hasta su evaluación final.

En cuanto a las actividades de aprendizaje, se analizaron 116 *frases* de las cuales se coincidió en 104 (89.65%) de ellas donde 74 cumplían con el patrón y 30 fueron enviadas a revisión. A pesar de que en las competencias se coincidió en que 6 serían enviadas a revisión y 13 estaban correctamente redactadas, la correlación entre estos dos grupos dio como resultado un 74.56% de efectividad al evaluarse un total de 114 relaciones donde 53 (46.49%) indicaron que existía una correspondencia entre la competencia y las actividades, mientras que 32 (28.07%) fueron enviadas directamente a revisión.

En cuanto al match entre las actividades de aprendizaje y las evidencias de aprendizaje, a pesar de que se relacionaron 116 actividades contra 65 evidencias se obtuvo una efectividad mayor al 95% lo cual se debe a los 3 análisis en conjunto que permitieron obtener un único resultado, donde solo 7 (3.86%) de las 181 relaciones encontradas en ambos sentidos (¿La actividad tiene una evidencia asociada? ¿La evidencia tiene una actividad asociada?), resultado que aprueba el uso de este método para relacionar ambos grupos.

Tabla 36-Resultados finales

Actividad	Total	Igual	Diferente	Bien	Mal	Porcentaje
Competencia	19	19	0	13	6	100
Actividades de aprendizaje	116	104	12	74	30	89.65
Actividades de enseñanza	91	86	5	55	31	94.50
Evidencias de aprendizaje	65	61	4	40	21	93.84
Instrumentos de evaluación	65	65	0	63	2	100
Match Competencia-Aprendizaje	114	85	29	53	32	74.56
Match Aprendizaje-Evidencia	116	112	4	82	30	96.55
Match Evidencia-Aprendizaje	65	62	3	58	4	95.38
Match Evidencia-Instrumentos	65	65	0	59	6	100
Total	715	658	57	496	162	92.02

Como conclusión, a pesar de que el modelo está basado en conocimiento de expertos y el algoritmo permite la extracción de información y características del texto, el conocimiento empírico para relacionar los temas y actividades favorecen el análisis por parte del evaluador, por lo que en procesos como la búsqueda de relación entre la competencia y las actividades de aprendizaje su resultado es superior al del módulo. No obstante, el método

permitió la identificación de los elementos claves de la ID manteniendo resultados semejantes a los del evaluador con un tiempo promedio de 15 segundos por unidad evaluada, mientras a un evaluador le toma aproximadamente 10 minutos realizar la misma tarea lo cual fundamenta su utilización como apoyo mas no siendo considerado un sustituto a su labor.

Capítulo 6-Conclusiones

A partir del desarrollo de esta investigación, se ha probado un algoritmo de PLN que en conjunto con un modelo de conocimiento han permitido el análisis de ID de docentes de educación superior. Basándose en el conocimiento de expertos del área, fue posible la codificación del algoritmo para la detección de patrones en los escritos en formato libre, apoyado por técnicas de extracción y recuperación de información.

De las pruebas y el análisis de los resultados se han considerado las siguientes conclusiones, recomendaciones y propuestas para trabajo a futuro.

6.1.-Conclusiones

La subjetividad sobre la interpretación de la técnica adecuada para la redacción de una ID, recae principalmente en el desconocimiento de la metodología a seguir para su elaboración en conjunto con la perspectiva propia del docente al considerar a su método como el mejor para la creación de una ID. A la par, los grandes volúmenes de información que son creados con las planeaciones de curso al inicio de cada semestre y el poco personal certificado para su evaluación, han mermado la retroalimentación que se otorga para corregir errores tanto de redacción como de coherencia.

Siendo esta una de las motivaciones, la presente investigación se centró en el diseño e implementación de un algoritmo basado en la metodología para la redacción de una ID de acuerdo con el Modelo EBC. Como la aplicación es específica, se desarrolló un modelo de conocimiento con los conceptos y relaciones del dominio en cuestión, identificando los patrones de las secciones que componen la ID, así como los elementos clave que son necesarios para cumplir con la metodología. La información que da respuesta a esas interrogantes fue recabada en base a consultas con expertos del área, lectura de documentos y el análisis de ID de años anteriores.

Como fruto de la investigación, se encontró la relación del proceso de aprendizaje con la Taxonomía de Bloom, pedagogo que lo divide en seis niveles que parten desde el conocimiento hasta la evaluación, por lo que se utilizó para clasificar los verbos y actividades de acuerdo con el dominio. Por lo tanto, gran parte del algoritmo se basa en el resultado del análisis de la competencia específica de cada unidad estructuradas de acuerdo con la taxonomía, lo que la convirtió en una herramienta ideal para el proceso.

Como resultado de estos análisis, se obtuvieron tablas con los elementos claves de una ID, así como las relaciones entre ellos normalizando el conocimiento a partir de la información disponible, la cual se modeló y se implementó computacionalmente utilizando el software Protégé. Ahora bien, la ontología contiene elementos extras que son predeterminados para cada materia y que puede ser encontrados en su temario, pero debido a que esa información no es relevante para el análisis que se pretendía y solo es una comprobación de existencia, se validó directamente en la plataforma y no fue incluido en el módulo de análisis.

Por otra parte, para el algoritmo de PLN se utilizaron principalmente el nivel de análisis morfológico y sintáctico con el fin de determinar las características gramaticales de las palabras y su papel estructural dentro de la *frase*, haciendo un preprocesamiento previo del texto en cuestión para evitar tener información no relevante. A pesar de tratarse de frases cortas, se optó por quitar las palabras vacías y normalizar el texto para abarcar la mayor cantidad de posibilidades que se pueden encontrar en una ID.

En cuanto a la comunicación entre el algoritmo de PLN y el modelo ontológico, se utilizó el lenguaje ontológico SPARQL para las consultas y un razonador deductivo para la inferencia de conocimiento. En conjunto, el algoritmo de PLN permitió la evaluación de ID de docentes de educación superior, analizando frases cortas correspondientes a las actividades plasmadas, así como las relaciones entre las actividades y las evidencias de aprendizaje para finalmente asociar las evidencias con los instrumentos que serán utilizados para otorgar la evaluación.

En las pruebas realizadas se encontraron deficiencias en la redacción de las actividades donde son omitidos elementos básicos como el verbo para identificar el dominio o el tema que será tratado con dicha actividad. A pesar de que la mayor parte de los análisis se realizaron a ID redactadas por docentes que aún no han llevado el curso de capacitación y formación de la plataforma, en el caso de la prueba piloto quedó demostrado que es posible que un docente redacte su ID en base a la metodología del modelo EBC y el algoritmo sea capaz de detectar si cumple o no con el patrón.

Como principal contribución, está el diseño del modelo de conocimiento aplicado a ID de educación superior, así como el algoritmo de PLN para la extracción de información con el fin de encontrar los patrones de redacción. Los resultados obtenidos denotan el cumplimiento de los objetivos planteados al inicio de la investigación y la factibilidad de la aplicación del algoritmo para otorgar una evaluación al momento, donde se obtuvo una eficiencia del 92% evaluando a ID del área superior tecnológica. Por otra parte, el tiempo que este requiere para analizar una ID completa es de aproximadamente 25 segundos, a diferencia del tiempo que requiere un evaluador certificado que puede tardar más de 30 minutos para la misma labor, por lo que el módulo puede servir de igual forma para aminorar su carga de trabajo haciéndoles partícipe únicamente de la evaluación final.

6.2.-Trabajo a Futuro

El algoritmo base se implementó como un módulo perteneciente a una plataforma de capacitación y formación docente, sin embargo, las pruebas se realizaron recibiendo los datos en formato JSON en el esquema que la plataforma utilizará una vez que esta esté en funcionamiento. Por este motivo, es necesario realizar pruebas piloto en la comunicación con

esta, así como evaluar instrumentaciones de docentes que ya han llevado el curso de formación y capacitación de la plataforma con el fin de comparar si los resultados de las pruebas a ID de años anteriores, a las pruebas piloto realizadas y a las obtenidas una vez que se ha tomado el curso permiten confirmar que la eficiencia del algoritmo aumentará una vez que se siga la metodología propuesta.

En cuanto al modelo de conocimiento, los datos predeterminados tanto de las actividades como de los verbos de la taxonomía de Bloom son pocos, siendo esta una limitante en el proceso. Para aminorar esta insuficiencia, es necesario una ampliación de los conceptos que contiene la ontología para tener una mayor cobertura, en especial en aquellas materias cuyas actividades son muy específicas y es necesario un vocabulario del dominio delimitado para que el algoritmo pueda hacer la detección en base a la información almacenada.

En cuanto al algoritmo, en uno de los procesos se utilizó una técnica de agrupamiento basada en aprendizaje no supervisado. A futuro, a partir de una colección de ID redactadas siguiendo la metodología de la plataforma, es factible el utilizar aprendizaje máquina para la detección de patrones entre ellas, extrayendo las características y relaciones que no han sido consideradas en el algoritmo actual debido a la cantidad limitada de ID con las que se contó para las pruebas. También, este apartado ayudaría a mejorar la detección de la relación entre la competencia y las actividades de aprendizaje.

Finalmente, las librerías de PLN evolucionan año con año, por lo que, si se retoma la investigación, las técnicas mejoradas y la precisión de las librerías para el análisis de textos en formato libre para el idioma español pueden permitir una mejor identificación de los patrones de los elementos que componen una ID. Además, al tener una investigación previa, permitiría tener un punto de partida sobre el cual medir los resultados que se obtengan de dicha investigación.

Bibliografía

- [1] F. I. Lanuza Gámez, M. Rizo Rodríguez, and L. E. Saavedra Torres, “Uso y aplicación de las TIC en el proceso de enseñanza-aprendizaje,” *Revista Científica de FAREM-Estelí*, pp. 16–30, 2018, [Online].
Available: <https://www.lamjol.info/index.php/FAREM/article/view/5667/5379>.
- [2] B. Zempoalteca Durán, J. Barragán López, J. González Martínez, and T. Guzmán Flores, “Formación en TIC y competencia digital en la docencia en instituciones de educación superior públicas,” *Apertura*, vol. 9, pp. 80–96, 2017, [Online].
Available: <http://www.scielo.org.mx/pdf/apertura/v9n1/2007-1094-apertura-9-01-00080.pdf>.
- [3] K. v. Pozos Pérez and J. Tejada Fernández, “Competencias Digitales en Docentes de Educación Superior: Niveles de Dominio y Necesidades Formativas,” *Revista Digital de Investigación en Docencia Universitaria*, vol. 12, no. 2, pp. 59–87, 2018, doi: 10.19083/ridu.2018.712.
- [4] R. del C. Martínez Rodríguez, L. Benítez Corona, and J. A. Vásquez Mora, “La instrumentación didáctica en transición de una educación tradicional a una basada en competencias,” *ANFEI Digital*, vol. 1, pp. 1–8, 2014.
- [5] J. Ángel. Vera Noriega, L. Elisa. Torres Moran, and E. Emmanuel. Martínez García, “Evaluación de competencias básicas en TIC en docentes de educación superior en

- México,” *Pixel-Bit. Revista de Medios y Educación*, no. 44, pp. 143–155, 2014, [Online]. Available: <http://www.redalyc.org/articulo.oa?id=36829340010>.
- [6] S. Arroyo Martínez, “Innovación en el diseño de modelos educativos basados en competencia en las instituciones de educación superior en México,” *Revista de Educación Superior*, vol. 2, no. 5, pp. 20–31, 2018, [Online]. Available: https://www.ecorfan.org/republicofperu/research_journals/Revista_de_Educacion_Superior/vol2num5/Revista_de_Educación_Superior_V2_N5_3.pdf.
- [7] H. Torres Ríos, I. H. Larios Hajar, and A. C. Medina Alcázar, “Neuroaprendizaje, actividades de enseñanza, actividades de aprendizaje e instrumentos de evaluación del aprendizaje,” *Debates en evaluación y currículum.*, vol. 24 de marzo, no. 4, 2018, [Online]. Available: <https://posgradoeducacionuatx.org/pdf2018/A220.pdf>.
- [8] R. J. Feo Mora, “Orientaciones básicas para el diseño de estrategias didácticas,” *Tendencias Pedagógicas*, no. 16, pp. 220–236, 2010, [Online]. Available: https://www.researchgate.net/publication/48523396_Orientaciones_basicas_para_el_diseno_de_estrategias_didacticas.
- [9] M. Suárez González, “Influencia del relativismo en la instrumentación didáctica y el desempeño escolar,” *RECIE. Revista Electrónica Científica de Investigación Educativa*, vol. 4, pp. 179–189, 2018, [Online]. Available: <https://www.rediech.org/ojs/2017/index.php/recie/article/view/318>.
- [10] J. A. Durango Hernández and Y. S. Pascuas Rengifo, “Los Sistemas Tutores Inteligentes Y Su Aplicabilidad En La Educación,” *Horizontes Pedagógicos*, vol. 17, no. 2, pp. 104–116, 2016, [Online]. Available: <https://revistas.iberamericana.edu.co/index.php/rhpedagogicos/article/view/826/772>.
- [11] C. C. Garnica, E. A. Sierra, B. B. Martínez, P. C. Márquez, J. Luis, and G. Cruz, “Elaboración de una Ontología para apoyar el diseño de secuencias didácticas basadas en competencias en la práctica del docente de educación media superior,” *Research in*

- Computing Science*, vol. 99, pp. 115–126, 2015, [Online]. Available: <https://pdfs.semanticscholar.org/9dc7/b5f410365f45fc9dc03784ce15a2c76f7f4c.pdf>.
- [12] L. R. T. Muñoz García, M. G. Gómez Zermeño, and L. Y. Alemán de la Garza, “Uso de la plataforma educativa Moodle en los procesos de capacitación de maestros de Educación Indígena en Jalisco, México,” *Revista del Instituto de Estudios de Educación Universidad del Norte*, pp. 42–73, 2016, doi: <http://dx.doi.org/10.14482/zp.24.8719>.
- [13] A. P. S. Cruz, “Principales Cifras del Sistema Educativo Nacional 2018-2019,” México, 2019. [Online]. Available: https://www.planeacion.sep.gob.mx/Doc/estadistica_e_indicadores/principales_cifras/principales_cifras_2018_2019_bolsillo.pdf.
- [14] A. F. Jaramillo, J. F. Duitama, J. F. Vélez, and R. I. Mazo, “Búsqueda de objetos de aprendizaje a partir de Objetivos instruccionales escritos en lenguaje natural,” *Tecnológicas*, no. 19, p. 185, 2007, doi: 10.22430/22565337.510.
- [15] C. A. Sánchez Muñoz, “El uso de la Tecnología de la Información y Comunicación como práctica docente en un modelo basado en competencias,” *Revista Iberoamericana de Producción Académica y Gestión Educativa*, vol. 5, 2018, [Online]. Available: <https://www.pag.org.mx/index.php/PAG/article/view/747>.
- [16] M. del C. de Luna Flores and J. J. Díaz Perera, “Modelo Educativo Mexicano: un reto para los docentes de matemáticas,” *Revista Electrónica sobre tecnología, Educación y Sociedad*, vol. 4, no. 7, 2017, [Online]. Available: <http://www.ctes.org.mx/index.php/ctes/article/viewFile/629/667>.
- [17] H. Solórzano Zamora and H. H. Caballero Vera, “Innovación metodológica para elevar el nivel de aprendizaje de la Química,” *Revista Didasc@lia: Didáctica y Educación*, vol. 10, pp. 161–176, 2019, [Online]. Available: <https://dialnet.unirioja.es/servlet/articulo?codigo=7242008>.

- [18] C. Cázares Sánchez, “Una intervención educativa con el sistema B-learning para la enseñanza del derecho, en la facultad de derecho, UNAM,” *Letras jurídicas*, pp. 17–32, 2016, [Online]. Available: <http://www.letrasjuridicas.com.mx/wp-content/uploads/2016/08/R34-Art01.pdf>.
- [19] MsC. A. Hernández Fernández, “Modelo Ontológico de recuperación de información para la toma de decisiones en Gestión de Proyectos Departamento de Información Y Comunicación,” 2015.
- [20] M. F. Ortega Carbajal, J. S. Hernández Mosqueda, and S. Tobón Tobón, “Análisis documental de la gestión del conocimiento mediante la cartografía conceptual,” *Ra Ximhai*, vol. 11, no. 4, pp. 141–160, 2015, doi: 10.35197/rx.11.01.e2.2015.09.mo.
- [21] C. A. Torres Navarro and J. A. Córdova Neira, “Diseño de sistema experto para toma de decisiones de compra de materiales,” *Cuadernos de Administración*, vol. 30, no. 52, pp. 20–30, 2014, doi: 10.25100/cdea.v30i52.27.
- [22] R. Fuentes Covarrubias, F. C. A. Gerardo, J. A. Cortes Quiroz, and J. G. de Jesús Juárez, “Sistema basado en conocimiento para la predicción del clima para usos agrícolas,” *Revista de Cómputo Aplicado*, vol. 2, no. 8, pp. 1–11, 2018, [Online]. Available: <http://www.scielo.org.co/pdf/cuadm/v30n52/v30n52a03.pdf>.
- [23] D. Cedeño Moreno and M. Vargas Lombardo, “Ontología y Procesamiento de Lenguaje Natural,” *KnE Engineering*, vol. 3, no. 1, pp. 492–501, 2018, doi: 10.18502/keg.v3i1.1453.
- [24] J. A. Guzmán Luna, M. López Bonilla, and I. Durley Torres, “Metodologías y métodos para la construcción de ontologías,” *Scientia et Technica*, vol. XVII, pp. 133–140, 2012, [Online]. Available: https://www.researchgate.net/publication/281120676_Metodologias_y_metodos_para_la_construccion_de_ontologias_Methodologies_and_methods_for_building_ontologies.

- [25] M. A. Musen, “The protégé project: A Look Back and a Look Forward,” *A. Musen, Mark*, vol. 1, no. 4, pp. 4–12, 2015, doi: 10.1145/2757001.2757003.
- [26] L. Talamé, A. Cardoso, and M. Amor, “Comparación de herramientas de procesamiento de textos en español extraídos de una red social para Python,” *XX Simposio Argentino de Inteligencia Artificial (ASAI 2019)-JAIIO 48 (Salta)*, vol. 1, pp. 53–67, 2019, [Online]. Available: <http://sedici.unlp.edu.ar/handle/10915/87854>.
- [27] A. C. Fernández Orquín, “Aplicación de la Semántica Multidimensional, Alineamiento Léxico-Semántico y Distancias léxicas al mejoramiento de tareas intermedias del PLN,” 2014.
- [28] E. Perdomo Sánchez, J. Díaz Blanco, A. Ojeda González, and N. Amador Domínguez, “Análisis de los procesos de lematización y estemizado en lingüística computacional,” *X Conferencia Internacional lingüística 2017*, no. November, 2017, [Online]. Available: https://www.researchgate.net/publication/322364515_analisis_de_los_procesos_de_lematizacion_y_estemizado_en_linguistica_computacional.
- [29] M. B. Hernández and J. M. Gómez, “Aplicaciones de Procesamiento de Lenguaje Natural,” *Revista Politécnica*, vol. 32, no. 1, pp. 87–96, 2013, [Online]. Available: http://www.revistapolitecnica.epn.edu.ec/ojs2/index.php/revista_politecnica2/article/view/32.
- [30] P. Gamallo, J. C. Pichel, M. García, J. M. Abuín, and T. Fernández-Pena, “Análisis morfosintáctico y clasificación de entidades nombradas en un entorno Big Data,” *Procesamiento de Lenguaje Natural*, vol. 53, pp. 17–24, 2014.
- [31] D. Reinsel, J. Gantz, and J. Rydning, “The Digitization of the World - From Edge to Core,” *IDC White Paper*, no. November, pp. 1–27, 2018, [Online]. Available: <https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf>.

- [32] I. Amón and C. Jiménez, “Funciones de Similitud sobre Cadenas de Texto: Una Comparación Basada en la Naturaleza de los Datos,” *International Conference on Information Resources Management*, p. 13, 2010, [Online]. Available: <https://core.ac.uk/download/pdf/11052347.pdf>.
- [33] M. Á. Álvarez Carmona, “Detección de similitud en textos cortos considerando traslape, orden y relación semántica de palabras,” 2014.
- [34] S. Bird, E. Klein, and E. Loper, *Natural Language Processing with Python*, 1st ed. O’Reilly Media, Inc., 2009.
- [35] L. Padró and E. Stanilovsky, “FreeLing 3 . 0 : Towards Wider Multilinguality,” *Proceedings of the Language Resources and Evaluation Conference (LREC 2012) ELRA*, pp. 2473–2479, 2012, [Online]. Available: <http://nlp.lsi.upc.edu/publications/papers/padro12.pdf>.
- [36] C. Manning, M. Surdeanu, J. Bauer, J. Finkel, S. Bethard, and D. McClosky, “The Stanford CoreNLP Natural Language Processing Toolkit,” *Association for Computational Linguistics (ACL) System Demonstrations*, pp. 55–60, 2014, doi: 10.3115/v1/p14-5010.
- [37] M. Honnibal and M. Johnson, *An improved non-monotonic transition system for dependency parsing*, no. September. Lisbon, Portugal: Association for Computational Linguistics, 2015.
- [38] T. Demeester, T. Rocktäschel, and S. Riedel, “Lifted rule injection for relation embeddings,” *EMNLP 2016 - Conference on Empirical Methods in Natural Language Processing, Proceedings*, pp. 1389–1399, 2016, doi: 10.18653/v1/d16-1146.
- [39] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” no. Mlm, 2018, [Online]. Available: <http://arxiv.org/abs/1810.04805>.

- [40] H. R. Gonzáles Argote and U. A. Ticona Gonzáles, “Clustering, mediterraneidad y comercio internacional: aplicación empírica de los algoritmos Partitioning Around Medoids y K-means,” *Revista Latinoamericana de Desarrollo Económico*, no. 32, pp. 96–130, 2019, doi: 10.35319/lajed.201932400.
- [41] G. Arroyo Delgado, “Metodología para elaborar el diseño de la enseñanza por competencia profesionales”, *Memorias del congreso CIIDET 2016*, México 2016.

Anexos

Anexo 1-Glosario de conceptos y su descripción

Nombre	Descripción
Actividades de Aprendizaje (AA)	Actividades realizadas por los estudiantes para reforzar los conocimientos proporcionados por el docente. Están divididas en las actividades predefinidas de acuerdo con la competencia seleccionada y al dominio al cual pertenecen en base a la Taxonomía de Bloom.
AA Predefinidas (AAP)	
AAP de análisis	
AAP de aplicación	
AAP de comprensión	
AAP de conocimiento	
AAP de evaluación	
AAP de síntesis	
AA seleccionadas	
Actividades de Enseñanza (AE)	Las actividades que el docente llevará a cabo para que el estudiante desarrolle con éxito la o las competencias genéricas y específicas establecidas para el tema. Están divididas en las actividades predefinidas de acuerdo con la competencia seleccionada y al dominio al cual pertenecen en base a la Taxonomía de Bloom.
AE Predefinidas (AEP)	
AEP de análisis	
AEP de aplicación	
AEP de comprensión	
AEP de conocimiento	
AEP de evaluación	
AEP de síntesis	
AEP seleccionadas	
Análisis de competencias específicas	Análisis de cada una de las competencias específicas a partir de los cinco objetos (temas y subtemas, actividades de aprendizaje, actividades de enseñanza, evidencias de aprendizaje e instrumentos de evaluación).
Apoyos didácticos	Se considera cualquier material que se ha elaborado para el estudiante con la finalidad de guiar los aprendizajes, proporcionar información, ejercitar sus habilidades, motivar e impulsar el interés y proporcionar un entorno de expresión.
Asignatura	Conjunto de objetivos de aprendizaje, contenidos y actividades que forma parte del plan de estudio de una carrera.
Competencias	Conocimientos, actitudes y habilidades que el alumno desarrollará durante el proceso de aprendizaje.
Competencias específicas	Son las competencias útiles para un ámbito disciplinar concreto, por lo tanto, no son transferibles a otros ámbitos.

Competencias específicas del docente	Conocimientos, habilidades y actitudes a desarrollar por el estudiante que el docente considera que es necesario implementar aun cuando las competencias de la asignatura son predeterminadas.
Competencias genéricas	Se refiere a las competencias indispensables en la formación de personas autónomas, tanto en el ámbito del aprendizaje como individual y social.
Competencias genéricas del docente	Conocimientos, habilidades y actitudes a desarrollar por el estudiante que el docente considera que es necesario implementar aun cuando las competencias de la asignatura son predeterminadas.
Competencias previas	Conocimientos, habilidades y actitudes que el estudiante debe de tener para poder cursar la materia.
Desarrollo de competencias genéricas	Con base en las actividades de aprendizaje establecidas en los temas, analizarlas en su conjunto y establecer que competencias genéricas se están desarrollando con dichas actividades.
Instrumentales	Se refieren a capacidades cognitivas, metodológicas, destrezas tecnológicas y destrezas lingüísticas a desarrollar por el estudiante durante un curso.
Instrumentales seleccionados	Competencias instrumentales que el docente ha seleccionado para desarrollar durante el curso.
Interpersonales	Son las competencias que facilitan el proceso de desarrollo humano y la interacción social del estudiante.
Interpersonales seleccionadas	Competencias interpersonales que el docente ha seleccionado para desarrollar durante el curso.
Sistemáticas	Tipo de competencias que permiten al estudiante comprender las relaciones existentes entre diferentes disciplinas o campos del conocimiento.
Sistemáticas seleccionadas	Competencias sistemáticas que el docente ha seleccionado para desarrollar durante el curso.
Evaluación	Proceso que permite determinar el grado de asimilación de los contenidos de una asignatura. En términos de enfoque por competencias se define como el proceso que permite determinar el grado de desarrollo de una competencia, a partir de criterios de desempeño.
Evaluación de la competencia	Instrumentos de evaluación de tipo formativo tales como las rúbricas o exámenes, los cuáles pueden ser considerados para evaluar el cumplimiento de las competencias planteadas.

Evaluación de evidencias	Contiene la instrumentación de evaluación en conjunto con una relación para evaluar las evidencias de aprendizaje que se han seleccionado.
Evidencias de Aprendizaje (EA)	<p>Conjunto de pruebas que demuestran los resultados del proceso de aprendizaje</p> <p>Están divididas en las evidencias predefinidas de acuerdo con la competencia seleccionada y al dominio al cual pertenecen de acuerdo con la Taxonomía de Bloom.</p>
EA predefinidas (EAP)	
EAP de análisis	
EAP de aplicación	
EAP de comprensión	
EAP de conocimiento	
EAP de evaluación	
EAP de síntesis	
EA seleccionadas	
Fuentes de información	Se considera a todos los recursos que contienen datos formales, informales, escritos, audio, imágenes, multimedia, que son utilizados por el docente para impartir la materia o por el estudiante para el logro de su aprendizaje.
Horas teórico practica	Duración del curso en horas divididas en actividades de carácter teórico y actividades de carácter práctico como talleres o laboratorios.
Indicadores de alcance	Dato que sirve para determinar el nivel de desempeño del estudiante, en relación con los objetivos de la materia o el logro de una competencia.
Instrumentos de Evaluación (IE)	<p>Herramientas que usará el profesor para obtener evidencias del desempeño del estudiante en el proceso de enseñanza-aprendizaje.</p> <p>Están divididos en los instrumentos predefinidos de acuerdo con la competencia seleccionada y al dominio al cual pertenecen de acuerdo con la taxonomía de Bloom.</p>
IE Predefinidos (IEP)	
IEP de análisis	
IEP de aplicación	
IEP de comprensión	
IEP de conocimiento	
IEP de evaluación	
IEP de síntesis	
IE seleccionados	
Matriz de evaluación	Tabla donde se establecen los criterios a evaluar en una tarea, así como la puntuación para cada uno de ellos.
Nivel de desempeño	Clasificación por categorías del logro de un estudiante, en el desarrollo de una competencia, de acuerdo con el desempeño que haya tenido en el curso.
Temas y subtemas	Partes en las que se divide el contenido de las unidades de un curso.

Tipo de evaluación	Forma de evaluación que se aplica durante el proceso de enseñanza y aprendizaje. Puede ser diagnóstica, formativa o sumativa.
Diagnostica	
Formativa	
Sumativa	
Valoración numérica	Escala establecida para asignar a un valor absoluto, en términos numéricos, al desempeño del estudiante en el desarrollo de una competencia.
Verbos	Acciones relacionadas con los objetivos de aprendizaje o la acción a realizar por el estudiante en el desarrollo de una competencia.
Verbos de análisis	Se considera una de las habilidades cognitivas superiores a desarrollar por el estudiante. Con ellos, se busca descomponer el problema de tal forma que sea posible analizarlo por partes y relacionar cada una de esas partes para poder llegar a una conclusión.
Verbos de aplicación	Habilidad cognitiva que hace referencia al uso práctico del conocimiento construido por el estudiante. Se refiere a usar el conocimiento en una nueva situación.
Verbos de comprensión	Este dominio se caracteriza por otorgarle al alumno la capacidad de comprender la información y expresarla con palabras propias.
Verbos de conocimiento	Hacen referencia a la capacidad del alumno por comprender por medio de la razón hechos, métodos, teorías, entre otros aspectos.
Verbos de evaluación	Estos verbos utilizan el juicio para analizar y posteriormente determinar criterios que permitan determinar el propósito de la actividad a realizar.
Verbos de síntesis	Habilidad del pensamiento que se define como la capacidad de un individuo para organizar y relacionar varios fragmentos de conocimiento en un todo coherente.

Anexo 2-Glosario de individuos y sus sinónimos

Individuo	sinónimos	
Abstraer	Abstracción	
Acordar	-	
Actividad focal inductora	-	
Agrupar	Agrupación	
Aislar	-	
Analizar	Análisis	
Analogía	-	
Aplicar	Aplicación	
Apoyar	-	
Apreciar	Apreciación	
Aprendizaje basado en análisis	Análisis descriptivo Análisis explicativo Análisis evaluativo Análisis cuantitativo Análisis cualitativo Análisis dato	Análisis video Descriptivo Explicativo Evaluativo Cualitativo Cuantitativo
Aprendizaje basado en problemas	Problema	
Aprendizaje basado en proyectos	Proyecto	
Aprovechar	-	
Argumentar	-	
Arreglar	-	
Artículo revista	Texto académico	
Asociar	Asociación	
Autoevaluación	-	
Búsqueda de Definición	-	
Calcular	-	
Calificar	Calificación	
Cambiar	-	
Catalogar	-	
Categorizar	Categorización	
Causa efecto	-	
Citar	-	
Clase práctica	Laboratorio Experimentación	Ejercicio Ejemplo
Clasificar	Clasificación	
Codificar	Codificación	
Coevaluación	-	
Combinar	Combinación	
Comparar	Comparación	
Compilar	Compilación	

Completar	-	
Componer		
Comprender	Comprensión	
Comprobar	Comprobación	
Computar	Computación	
Comunicación de forma escrita	Comunicación escrita texto	escrito
Comunicación oral	-	
Concebir	-	
Concepto	-	
Conceptuar	-	
Concluir	Conclusión	
Concretar		
Conducir lectura comentada	Reporte	Lectura
Conducir taller	taller	
Conectar	-	
Conocer	-	
Construir	Construcción	
Contrastar	-	
Convencer	-	
Convertir	-	
Crear	Creación	
Criticar	-	
Cuadro comparativo	Tabla comparativa Estudio comparativo	Tabla descriptiva
Cuadro Sinóptico	-	
Cuento	Historieta	Comic
Cuestionar	preguntar	
Cuestionario	-	
Debate	Debate Debatir Argumentar	Propuesta argumentada Opinar Opinión
Decidir	-	
Deducir	Deducción	
Defender	-	
Definir	Definición	
Delimitar	Delimitación	
Delinear	-	
Demostrar	Demostración	
Desarrollar	-	
Descodificar	Descodificación	
Descomponer	Descomposición	
Describir	Descripción	

Descubrir	-	
Desglosar	-	
Designar	Designación	
Desmenuzar	-	
Despejar	-	
Destacar	-	
Detallar	-	
Detectar	Detección	
Determinar	-	
Diagrama de árbol	-	
Diagrama de correlación	-	
Diagrama de flujo	-	
Diagrama de llaves	-	
Diagrama de Venn	-	
Diagrama radial	-	
Diagramar	-	
Dibujo	Dibujar Ilustración Fotografía	Bosquejo Collage Boceto
Diferenciar	Diferenciación	
Dirigir	-	
Discriminar	Discriminación	
Discurso explicativo	Discurso Expositivo	Discurso
Discusión de Casos	Discusión en grupo Panel de discusión	Discutir Discusión
Diseñar	Diseño	
Distinguir	Distinción	
Distribuir	Distribución	
Dividir	División	
Dramatizar	Dramatización	
Ejemplificar	Ejemplificación	
Ejemplo	Ejemplo	
Ejercicios	-	
Elaborar	Elaboración	
Elegir	Elección	
Eliminar	Eliminación	
Emplear	-	
Encontrar	-	
Enlistar	-	
Ensamblar	-	
Ensayo	-	
Entrevista semiestructurada	Entrevista libre	
Enumerar	Enumeración	
Enunciar	Enunciación	

Equipos	Pareja Dupla	Grupo Grupal
Erigir	-	
Esbozar	-	
Escala Estimativa	-	
Escoger	-	
Escribir	-	
Especificar	Especificación	
Esquematizar	Esquemmatización	
Establecer	-	
Estimar	Estimación	
Estructurar	Estructuración	
Estudio de caso	-	
Estudio dirigido	Exposición	Presentación
Evaluar	Evaluación	
Examen	-	
Examinar	-	
Experimentar	Experimentación	Experimento
Explicar	Explicación	
Exponer	Exposición Presentación	Diapositivas Power point
Expresar	-	
Extender	-	
Extrapolar	Extrapolación	
Ficha De lectura	-	
Formular	Formulación	
Fraccionar	-	
Fundamentar	Fundamentación	
Generalizar	Generalización	
Generar	Generación	
Glosario	-	
Grafica	-	
Guía de observación	-	
Heteroevaluación	-	
Heurística uve de Gowin	-	
Hipótesis	-	
Hoja de calculo	Excel	Base de dato
Idear	-	
Identificar	Identificación	
Ilustrar	Ilustración	
Indagación	Indagación	
Indicar	Indicación	
Inducir	Inducción	
Inferir	-	

Infografía	-	
Informar	Informe	
Inspeccionar	-	
Integrar	Integración	
Interpolar	Interpolación	
Interpretar	Interpretación	
Inventar	-	
Inventariar	-	
Investigación	Investigación documental Investigación Internet Investigación Web	Reporte de investigación Investigar
Jerarquizar	Jerarquización	
Justificar	Justificación	
Juzgar	-	
Línea de tiempo	-	
Lista de cotejo	-	
Listar	-	
Localizar	-	
Manejar	-	
Manipular	-	
Mapa conceptual	-	
Mapa mental	-	
Maqueta	-	
Marcar	-	
Medir	-	
Memorizar	Memorización	
Mencionar	-	
Metodología	Instrumento	
Modelo	Modelar	
Modificar	-	
Mostrar	-	
Narrar	Narración	
Nombrar	-	
Obtener	-	
Omitir	-	
Operar	Operación	
Ordenar	-	
Organizador grafico QQQ	Gráfico QQQ	
Organizar	Organización	
Parafrasear	-	
Pedir	-	
Película	Video	Audiovisual
Planear	Planeación	
Plantear	-	

Podcast	-	
Portafolio de evidencias	-	
Poster	Cartel	
Practica de laboratorio	-	
Practicar	-	
Precisar	-	
Predecir	Predicción	
Preguntas	Pregunta Intercalada	
Preparar	Preparación	
Probar	-	
Producir	Producción	
Programar	Programación	
Pronosticar	-	
Proponer	Proposición	
Propuesta	-	
Proyectar	Proyección	
Radio	-	
Reacomodar	Reacomodación	
Reafirmar	Reafirmación	
Realizar	Realización	
Recoger	-	
Recolección de datos	Recolección de muestra	Muestra
Recolectar	Recolección	
Recomendar	Recomendación	
Reconocer	Reconocimiento	
Reconstruir	Reconstrucción	
Recordar	-	
Redactar	-	
Reescribir	Reescritura	
Referir	-	
Reflexionar	Reflexión	
Refutar	Refutación	
Registrar	-	
Registro Descriptivo	Registro Anecdótico	
Relacionar	-	
Relatar	-	
Reordenar	-	
Reorganizar	Reorganización	
Reparar	Reparación	
Repasar	-	
Repetir	Repetición	
Reportar	Reporte	
Representar	Representación	
Reproducir	Reproducción	

Reseña	-	
Resolver	Resolución	
Resumir	Resumen	
Reunir	Reunión	
Revisar	-	
Rotular	Rotulación	
Rúbrica	-	
Seccionar	Sección	
Seguir	-	
Seleccionar	-	
Señalar	Señalización	
Separar	Separación	
Seriar	Serialización	
Silogismo	-	
Simular	Simulación	
Síntesis	-	
Sintetizar	-	
Sostener	-	
Subdividir	Subdivisión	
Subrayar	-	
Subsistir	-	
Sumar	-	
Sustentar	Sustentación	
Tabular	Tabulación	
Tasar	-	
Teatro	-	
Teorizar	-	
Test	Examen	
Traducir	Traducción	
Transcribir	Transcripción	
Transferir	-	
Transformar	Transformación	
Trazar	-	
Usar	Uso	
Uso de Tic	TIC Tecnologías Información Comunicación	Computadora Laptop PC
Utilizar	Utilización	
Valorar	Valoración	
Valuar	-	
Verificar	Verificación	
Visita Industrial	Visita Fabrica	
Viaje usuario	-	

Anexo 3-Taxonomía de conceptos

- ✓ Actividades de aprendizaje
 - Actividades de aprendizaje predefinidas
 - Actividades de aprendizaje predefinidas de análisis
 - Actividades de aprendizaje predefinidas de aplicación
 - Actividades de aprendizaje predefinidas de comprensión
 - Actividades de aprendizaje predefinidas de conocimiento
 - Actividades de aprendizaje predefinidas de evaluación
 - Actividades de aprendizaje predefinidas de síntesis
 - Actividades de aprendizaje seleccionadas
- ✓ Actividades de enseñanza
 - Actividades de enseñanza predefinidas
 - Actividades de enseñanza predefinidas de análisis
 - Actividades de enseñanza predefinidas de aplicación
 - Actividades de enseñanza predefinidas de comprensión
 - Actividades de enseñanza predefinidas de conocimiento
 - Actividades de enseñanza predefinidas de evaluación
 - Actividades de enseñanza predefinidas de síntesis
 - Actividades de enseñanza seleccionadas
- ✓ Análisis de competencias específicas
- ✓ Apoyos didácticos
- ✓ Asignatura
- ✓ Competencias
 - Competencias específicas
 - Competencias específicas del docente
 - Competencias genéricas
 - Competencias genéricas del docente
 - Competencias previas
- ✓ Desarrollo de competencias genéricas
 - Instrumentales
 - Instrumentales seleccionados
 - Interpersonales
 - Interpersonales seleccionadas
 - Sistemáticas
 - Sistemáticas seleccionadas
- ✓ Evaluación
- ✓ Evaluación de la competencia
- ✓ Evidencias de aprendizaje
 - Evidencias de aprendizaje predefinidas
 - Evidencias de aprendizaje predefinidas de análisis

- Evidencias de aprendizaje predefinidas de aplicación
 - Evidencias de aprendizaje predefinidas de comprensión
 - Evidencias de aprendizaje predefinidas de conocimiento
 - Evidencias de aprendizaje predefinidas de evaluación
 - Evidencias de aprendizaje predefinidas de síntesis
 - Evidencias de aprendizaje seleccionadas
- ✓ Evaluación de evidencias
- ✓ Fuentes de información
- ✓ Horas teórico practica
- ✓ Indicadores de alcance
- ✓ Instrumentos de evaluación
 - Instrumentos de evaluación predefinidos
 - Instrumentos de evaluación predefinidos de análisis
 - Instrumentos de evaluación predefinidos de aplicación
 - Instrumentos de evaluación predefinidos de comprensión
 - Instrumentos de evaluación predefinidos de conocimiento
 - Instrumentos de evaluación predefinidos de evaluación
 - Instrumentos de evaluación predefinidos de síntesis
 - Instrumentos de evaluación seleccionados
- ✓ Matriz de evaluación
- ✓ Nivel de desempeño
- ✓ Temas y subtemas
- ✓ Tipo de evaluación
 - Diagnostica
 - Formativa
 - Sumativa
- ✓ Valoración numérica
- ✓ Verbos
 - Verbos de análisis
 - Verbos de aplicación
 - Verbos de comprensión
 - Verbos de conocimiento
 - Verbos de evaluación
 - Verbos de síntesis

Anexo 4-Propiedades de clases

DOMINIO	RELACIÓN	PROPIEDAD	TIPO
Análisis de competencias específicas	tiene	Descripción	String
		Diagnóstica	
		Fecha de final	
		Fecha de inicio	
		Formativa	
		Nombre	
		Número	
		Sumativa	
Asignatura	Tiene	Aportación	String
		Aula	
		Caracterización	
		Carrera	
		Ciclo escolar	
		Clave	
		Departamento	
		Descripción	
		Docente	
		Fecha	
		Grupo	
		Horas	
		Importancia	
		Intención didáctica	
		Nombre	
		Periodo	
		Plantel	
Relación con otras asignaturas			
Relación con competencias específicas			
Relación con temas			
Evidencias de aprendizaje	Tiene	Indicador	String
		Porcentaje	
		Valor del indicador	
Indicador de alcance	Tiene	valor	String

Anexo 5-Relaciones entre conceptos

Dominio	Relación	Rango
Análisis de competencias específicas	tiene	Actividades de aprendizaje seleccionadas
		Actividades de enseñanza seleccionadas
		Apoyos didácticos
		DCG instrumentales seleccionadas
		DCG interpersonales seleccionadas
		DCG sistemáticas seleccionadas
		Evaluación
		Fuentes de información
		Horas teórico-prácticas
		Temas y subtemas
Asignatura	tiene	Actividades de aprendizaje predefinidas
		Actividades de enseñanza predefinidas
		Análisis de competencias específicas
		Competencias específicas del docente
		Competencias genéricas
		Competencias genéricas de docentes
		Competencias previas
Evaluación	tiene	Evaluación de competencia
		Indicadores de alcance
		Matriz de evaluación
		Niveles de desempeño
Evidencia de aprendizaje	tiene	Instrumentos de evaluación seleccionados
Matriz de evaluación	Tiene	Evidencias de aprendizaje seleccionadas
Verbos de análisis	incluye	Actividades de aprendizaje de análisis
		Actividades de enseñanza de análisis
		Evidencias de aprendizaje de análisis
		Instrumentos de evaluación de análisis
Verbos de aplicación	incluye	Actividades de aprendizaje de aplicación
		Actividades de enseñanza de aplicación
		Evidencias de aprendizaje de aplicación
		Instrumentos de evaluación de aplicación
Verbos de comprensión	incluye	Actividades de aprendizaje de comprensión
		Actividades de enseñanza de comprensión
		Evidencias de aprendizaje de comprensión
		Instrumentos de evaluación de comprensión
Verbos de conocimiento	incluye	Actividades de aprendizaje de conocimiento

		Actividades de enseñanza de conocimiento
		Evidencias de aprendizaje de conocimiento
		Instrumentos de evaluación de conocimiento
Verbos de evaluación	incluye	Actividades de aprendizaje de evaluación
		Actividades de enseñanza de evaluación
		Evidencias de aprendizaje de evaluación
		Instrumentos de evaluación de evaluación
Verbos de evaluación	incluye	Actividades de aprendizaje de evaluación
		Actividades de enseñanza de evaluación
		Evidencias de aprendizaje de evaluación
		Instrumentos de evaluación de evaluación

Anexo 6-Relación entre evidencias e instrumentos

Evidencia de aprendizaje	R	L	E	P	X	G	A	G	H	I
Actividad focal inductora	1	1	1	1	0	0	0	0	0	0
Analogía	1	1	1	1	0	0	0	0	0	0
Aprendizaje basado en análisis	1	0	0	1	0	0	0	0	0	0
Aprendizaje basado en problemas	1	1	0	1	0	0	1	0	0	0
Aprendizaje basado en proyectos	1	1	0	1	1	0	0	0	0	0
Artículo de revista	1	1	1	1	0	0	0	0	0	0
Autoevaluación	1	1	1	1	0	0	1	0	0	0
Búsqueda definición	1	1	1	1	0	0	0	0	0	0
Causa efecto	1	1	0	1	0	0	0	0	0	0
Comunicación de forma escrita	1	1	1	1	0	0	1	0	0	0
Comunicación oral	1	0	1	1	0	0	1	0	0	1
Cuadro comparativo	1	1	1	1	0	0	0	0	0	0
Cuadro sinóptico	1	1	1	1	0	0	0	0	0	0
Concepto	1	1	1	1	1	0	0	0	0	0
Cuento	1	1	1	1	0	0	0	0	0	0
Cuestionario	1	1	1	1	0	1	1	0	0	0
Debate	1	1	0	1	0	0	0	0	0	1
Diagrama de árbol	1	1	1	1	0	0	0	0	0	0
Diagrama de correlación	1	1	1	1	0	0	0	0	0	0
Diagrama de flujo	1	1	1	1	0	0	0	0	0	0
Diagrama de llave	1	1	1	1	0	0	0	0	0	0
Diagrama de Venn	1	1	1	1	0	0	0	0	0	0
Diagrama radial	1	1	1	1	0	0	0	0	0	0
Dibujo	1	1	1	1	0	0	0	0	0	0
Discurso explicativo	1	1	1	1	0	0	0	0	0	1
Discusión de casos	1	1	1	1	0	0	0	0	0	1

Ejemplos	1	0	1	1	1	0	0	0	0	0
Ejercicios	1	0	1	1	1	0	0	0	0	0
Ensayo	1	1	1	1	0	0	1	0	0	0
Entrevista	1	1	1	1	0	1	1	0	0	0
Estudio de caso	1	1	0	1	0	0	1	0	0	0
Experimento	1	1	0	1	0	0	0	1	0	0
Glosario	1	1	1	1	0	0	0	0	0	0
Grafica	1	1	1	1	0	0	0	0	0	0
Heurística Uve de Gowin	1	1	1	1	0	0	0	0	0	0
Hipótesis	1	0	0	1	0	0	1	0	0	0
Hoja de calculo	1	1	1	1	0	0	0	0	0	0
Indagación	1	1	1	1	0	0	0	0	0	0
Infografía	1	1	1	1	0	0	0	0	0	0
Informe	1	1	1	1	0	0	0	0	0	0
Investigación	1	1	0	1	0	0	1	0	0	0
Línea de tiempo	1	1	0	1	0	0	0	0	0	0
Mapa conceptual	1	0	0	1	0	0	0	0	0	0
Mapa mental	1	0	0	1	0	0	0	0	0	0
Maqueta	1	1	1	1	0	0	0	0	0	0
Metodología	1	1	1	1	0	0	0	0	0	0
Modelo	1	0	1	1	1	0	0	0	0	0
Observación	1	1	1	1	0	0	0	0	1	0
Organizador grafico QQQ	0	1	1	1	0	0	0	0	0	0
Película	1	1	1	1	0	0	0	0	0	0
Poster	1	1	1	1	0	0	0	0	0	0
Practica	1	1	0	1	0	0	0	1	0	0
Pregunta	1	1	1	1	0	1	1	0	0	0
Presentación	1	0	1	1	0	0	0	0	0	0
Propuesta	1	1	1	1	0	0	0	0	0	0
Radio	1	1	1	1	0	0	0	0	0	0
Recolección de datos	1	1	0	1	0	0	0	0	0	0
Reflexionar	1	1	1	1	0	0	1	0	0	1
Reseña	1	0	1	1	0	0	0	0	0	0
Resumen	1	1	0	1	0	0	0	0	0	0
Simulación	1	1	1	1	0	0	0	1	0	0
Síntesis	1	0	1	1	0	0	0	0	0	0
Teatro	1	1	1	1	0	0	0	0	0	0
Trabajar en equipo	1	1	0	1	0	0	0	0	0	0
Uso de TIC	1	0	1	1	0	0	0	0	0	0
Viaje de usuario	1	0	1	1	0	0	0	0	0	0
Visita industrial	1	1	0	1	0	0	0	0	0	0