

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN
FACULTAD DE INGENIERÍA MECÁNICA Y ELÉCTRICA



**COMPARACIÓN DE REDES SEMÁNTICAS NATURALES
MEDIANTE UN MODELO DE RECUPERACIÓN DE INFORMACIÓN**

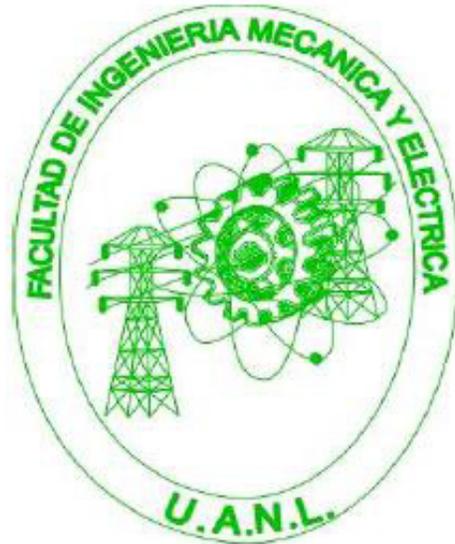
POR

AÍDA LUCINA GONZÁLEZ LARA

**EN OPCIÓN AL GRADO DE
DOCTOR EN INGENIERÍA CON ORIENTACIÓN EN
TECNOLOGÍAS DE LA INFORMACIÓN**

JULIO 2016

**UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN
FACULTAD DE INGENIERÍA MECÁNICA Y ELÉCTRICA
SUBDIRECCIÓN DE ESTUDIOS DE POSGRADO**



**COMPARACIÓN DE REDES SEMÁNTICAS NATURALES
MEDIANTE UN MODELO DE RECUPERACIÓN DE
INFORMACIÓN.**

**POR
AÍDA LUCINA GONZÁLEZ LARA**

**EN OPCIÓN AL GRADO DE:
DOCTOR EN INGENIERÍA CON ORIENTACIÓN EN
TECNOLOGÍAS DE LA INFORMACIÓN**

SAN NICOLÁS DE LOS GARZA, NUEVO LEÓN,

JULIO 2016

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN
FACULTAD DE INGENIERÍA MECÁNICA Y ELÉCTRICA
SUBDIRECCIÓN DE ESTUDIOS DE POSGRADO

Los miembros del comité de tesis recomendamos que la tesis "Comparación de Redes Semánticas Naturales mediante un Modelo de Recuperación de Información" realizada por la alumna Aida Lucina González Lara, con número de matrícula 330349, sea aceptada para su defensa como opción al grado de Doctor en Ingeniería con orientación en Tecnologías de la Información.

El Comité de Tesis



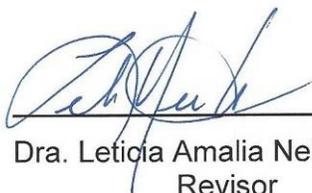
Dr. Francisco Torres Guerrero
Asesor



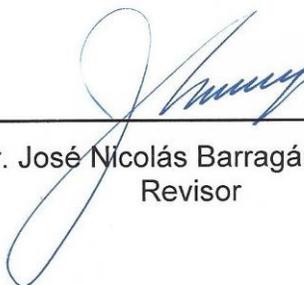
Dr. Jaime Arturo Castillo Elizondo
Revisor



Dr. Sergio Fernando Alcaraz Corona
Revisor



Dra. Leticia Amalia Neira Tovar
Revisor



Dr. José Nicolás Barragán Codina
Revisor

Vo. Bo.



Dr. Simón Martínez Martínez
Subdirector de Estudios de Posgrado

AGRADECIMIENTOS

Agradezco a Dios por darme fuerza y voluntad para lograr mis metas.

Mi profundo agradecimiento a mi familia, este trabajo fue posible gracias a la seguridad y estabilidad proporcionada por su amor y motivación.

De manera especial agradezco a mi asesor Dr. Francisco Torres por su supervisión y contribuciones en el desarrollo de esta tesis, así como a mis revisores Dr. Jaime A. Castillo, Dr. Sergio Alcaraz, Dra. Leticia Neira y Dr. Nicolás Barragán, por el tiempo dedicado y sus comentarios para mejorar este trabajo. Agradezco a mis profesores de doctorado por sus significativos conocimientos y experiencias compartidas.

Un franco agradecimiento al Dr. Jaime A. Castillo y al Dr. Simón Martínez por su confianza y apoyo permanente para lograr esta meta, así como al Dr. Moisés Hinojosa por su motivación para iniciar este doctorado, por su disposición y sus valiosos consejos.

Gracias a la Facultad de Ingeniería Mecánica y Eléctrica de la UANL, mi Alma Máter, en la que me he formado y desarrollado profesionalmente, por permitirme crecer y aprender cada día.

Agradezco sinceramente a todas las personas con las que he coincidido durante este tiempo: amigos, compañeros, tesisas, alumnos, que de diversas maneras me alentaron y apoyaron para a concluir este grado.

A mi familia,

el regalo más valioso que me ha dado la vida,

mi fuente constante e incondicional de apoyo y amor.

ÍNDICE GENERAL

RESUMEN	.X
INTRODUCCIÓN	1
1.1. Antecedentes	1
1.2. Justificación.....	4
1.3. Definición del Problema	5
1.4. Objetivos	6
1.4.1. Objetivo General.....	6
1.4.2. Objetivos Particulares.....	6
1.5. Hipótesis	7
1.6. Preguntas de investigación	7
1.7. Estructura de la tesis.....	8
MARCO TEÓRICO	9
2.1. Representación del conocimiento	9
2.1.1. Redes Semánticas.....	10
2.2. Redes Semánticas Naturales.....	12
2.3. Recuperación de Información	15

2.3.1. Sistemas de Recuperación de Información	16
2.3.2. Modelos de Recuperación de Información.....	17
2.3.2.1. Modelo Booleano	18
2.3.2.2. Modelo del Espacio Vectorial	19
2.3.2.3. Modelo Probabilístico	21
2.4. Similitud	22
2.4.1. Medidas de Similitud.....	23
2.4.1.1. Similitud Cosenoidal.....	23
2.4.1.2. Coeficiente de Jaccard.....	24
2.4.1.3. Coeficiente de Dice	24
2.4.2. Distancias como indicadores de disimilitud.....	24
2.4.3. Familia Minkowski de disimilitud	25
2.5. Visualización de Redes Semánticas Naturales	26
METODOLOGÍA	29
3.1. Descripción del estudio	29
3.2. Participantes	30
3.3. Especificación de técnica para realizar estudio de Redes Semánticas Naturales.....	31
3.4. Generación de Redes Semánticas Naturales	35
3.4.1. Aplicación de RSN a principiantes utilizando <i>RSNweb</i>	36
3.4.2. Limpieza de información y generación de RSN de principiantes.	38
3.5. Generación de RSN de expertos	39
3.5.1. Aplicación de RSN a expertos	40
3.6. Modelo de Espacio Vectorial en Redes Semánticas Naturales	41
3.6.1. Pre-procesamiento de la Información	42

3.7. Visualización de Redes Semánticas Naturales mediante nubes de etiquetas.	44
3.8. Comparación de RSN mediante el Modelo de Espacio Vectorial.....	48
3.8.1. Proceso de comparación de RSN mediante MEV	49
RESULTADOS Y DISCUSIÓN	53
4.1. Resultados de la generación de RSN para principiantes	53
4.1.1. Generación de Red Total de Principiantes	55
4.2. Resultados de la generación de Red de Expertos	56
4.3. Representación visual de la RSN generadas.....	59
4.3.1. Representación visual de RSN de principiantes	60
4.3.2. Representación visual de RSN de expertos	62
4.3.3. Representación visual de Redes Totales.....	64
4.4. Resultado de la comparación de Redes Semánticas Naturales mediante Modelo de Espacio Vectorial	65
4.4.1. Comparación entre RSN de principiantes.....	65
4.4.2. Análisis de resultados de comparación entre RSN de principiantes.....	67
4.4.3. Comparación entre RSN de expertos	69
4.4.4. Análisis de resultados de comparación entre RSN de expertos ..	72
4.4.5. Comparación entre Red Total de principiantes con RSN de expertos	74
4.4.6. Análisis de resultados de comparación entre Red Total de principiantes con RSN de expertos.....	77
4.4.7. Comparación entre RSN de principiantes con Red Total de expertos	78

4.4.8. Comparación entre Red Total de Principiantes con Red Total de Expertos.....	80
4.4.9. Resultados de la relación entre las medidas de similitud	83
CONCLUSIONES Y TRABAJO FUTURO	86
5.1. Conclusión del Modelo de Comparación de RSN	86
5.2. Conclusión de las distintas medidas de similitud aplicadas en la comparación de RSN mediante el MEV.....	89
5.3. Comentarios conclusivos	90
5.4. Trabajo Futuro.....	92
REFERENCIAS	93
ÍNDICE DE FIGURAS	104
ÍNDICE DE TABLAS	108
APÉNDICE A	111
APÉNDICE B	113
APÉNDICE C	115
APÉNDICE D	117

RESUMEN

Actualmente la técnica de Redes Semánticas Naturales (RSN) es usada ampliamente para la representación de significados, es de interés comparar redes de diferentes poblaciones acerca de una misma temática. En este trabajo se presenta un nuevo modelo de comparación entre RSN desde una perspectiva de Recuperación de Información para mejorar el alcance y optimizar la forma de comparar dos RSN, asimismo se genera un método de visualización de las RSN que permita representar a las redes así como a la similitud entre ellas. Para realizar esta investigación, se generaron diez RSN de la temática de Ecología con la participación de 163 individuos; se adaptó la definición del Modelo de Espacio Vectorial (MEV) para comparar las RSN, como resultado de la comparación se obtuvo el valor de tres medidas de similitud: cosenoidal, el coeficiente de Dice y el coeficiente de Jaccard; se calcularon coeficientes de correlación de Pearson para determinar la relación entre estas medidas y el resultado indica que la correlación que existe entre las medidas de similitud es positiva muy fuerte. Se generaron representaciones visuales de las redes mediante nubes de etiquetas en donde el peso de cada definidor está relacionado con el tamaño y color con el que se muestra, se identificó que esta visualización es una técnica útil para representar RSN así como para evaluar de manera cualitativa la similitud entre ellas. Se llevó a cabo el análisis de los resultados obtenidos en la comparación de manera cuantitativa y cualitativa el cual indica que sí es posible realizar la comparación entre RSN con este método aun cuando la cantidad de participantes en las redes sea distinta, lo que proporciona una nueva herramienta para la amplia variedad de campos en los que se usan la Redes Semánticas Naturales.

CAPÍTULO 1

INTRODUCCIÓN

1.1. Antecedentes

Las redes semánticas son estructuras gráficas utilizadas para la representación de conocimiento y conceptos en computadoras, sus aplicaciones incluyen tópicos tales como: visión artificial, comprensión del lenguaje natural y recuperación de información; las redes semánticas son un lenguaje visual del pensamiento [1]. Las redes semánticas no sólo representan información, sino que facilitan la recuperación de hechos relevantes debido a que organizan la información de tal manera que se facilite esta recuperación [2].

El uso de redes semánticas ha permitido dimensionar la representación del conocimiento en diferentes niveles en la aplicación de diversas áreas de investigación, existen distintos métodos para la construcción de redes semánticas, en el caso de las Redes Semánticas Naturales (RSN), esta es una técnica que se basa en la generación de definiciones conceptuales de participantes para, con dichas definiciones, construir una red semántica; la diferencia con otros métodos que producen redes semánticas, está en el origen de las definiciones con las que la red es construida, ya que en las RSN estas definiciones son proporcionadas por el propio participante [3]. En la Figura 1.1 se representa una RSN generada para la temática de Simulación de Sistemas.

SIMULACION 1

EVENTO		
F	Definidor	M
1	SUCESO	56
1	ACCION	43
3	RESULTADO	34
1	ACONTECIMI...	33
1	CLICK	30
4	SISTEMA	28
1	INTERACCI...	28
4	SIMULACION	27
1	FECHA	26
1	MOMENTO	25

SIMULADOR		
F	Definidor	M
4	SISTEMA	78
4	SIMULACION	69
3	VARIACION	46
1	PROGRAMA	46
2	PRUEBAS	36
2	MODELO	26
3	RESULTADO	25
2	EXPERIMENT...	24
2	SIMULAR	23
1	REPRODUCIR	22

NUMERO ALEATORIO		
F	Definidor	M
2	AZAR	99
2	RANDOM	67
3	VARIACION	66
2	NUMERO	49
1	CAMBIANTE	48
2	GENERADOR	35
2	DIFERENTE	32
1	SEMILLA	28
2	SIMULAR	17
1	METODO	16

MODELO		
F	Definidor	M
4	SIMULACION	64
1	PROBLEMA	37
2	REPRESENTA...	34
1	PROTOTIPO	34
3	VARIACION	33
4	SISTEMA	26
2	DATO	25
1	PROYECTO	24
1	MATEMATICO	19
1	COMPUTACIO...	18

SIMULACION		
F	Definidor	M
2	MODELO	53
4	SISTEMA	52
1	SIMULADOR	48
2	EXPERIMENT...	33
1	HIPOTESIS	31
1	PROCESO	29
2	PRUEBAS	28
3	RESULTADO	28
2	DATO	26
2	REPRESENTA...	26

VARIABLE ALEATORIA		
F	Definidor	M
2	AZAR	78
2	RANDOM	58
1	ALEATORIED...	40
2	GENERADOR	40
1	FUNCION	38
2	NUMERO	34
4	SIMULACION	33
2	DIFERENTE	27
1	IMPREDECIB...	26
1	CONSTANTE	25

Figura 1. 1. Representa una Red Semántica Natural generada para la temática de Simulación de Sistemas [4].

La representación y estructura del conocimiento son factores determinantes para el desempeño de tareas cognitivas, los adelantos en el estudio de la cognición han permitido, además de hacer hipótesis relacionadas con los procesos internos de adquirir y organizar estructuras de conocimiento, dar seguimiento a dichos procesos, así como generar visualizaciones de su evolución y estado [5].

Si se toma en cuenta que el conocimiento es relacional, evaluar este conocimiento involucra: la elicitación de juicios de relación, definir alguna representación de la estructura cognitiva basada en esos juicios de relación, evaluar la estructura cognitiva comparándola con una estructura referente y comparar las evaluaciones con otras medidas indicativas de conocimiento [6]. En la evaluación, al hablar de la comparación con una estructura referente, generalmente esta referencia es un experto, de ahí la importancia de métodos que permitan hacer comparaciones para determinar los cambios en el conocimiento [7].

Las tecnologías de la información (TI) han transformado la manera de procesar información hacia un efectivo tratamiento de la misma mediante el uso de herramientas para convertir dicha información en conocimiento [8]. Las ciencias computacionales ofrecen métodos de análisis de información, los cuales pueden ser aplicados en otras áreas de estudio con el objetivo de optimizar su desempeño.

La Recuperación de Información (*Information Retrieval*) es un campo de estudio que se relaciona con la estructura, el análisis, el almacenamiento y la búsqueda de información, tales acciones son requeridas para dar respuesta de manera eficiente a consultas de información que son realizadas por los usuarios [9] [10].

La finalidad más importante de la Recuperación de Información es mostrar a los usuarios aquella información que es relevante para ellos, como lo muestra la Figura 1.2 [11]; para lograr lo anterior los sistemas de recuperación de información aplican técnicas de evaluación con el objetivo de comparar diversos documentos con una consulta solicitada por el usuario, después de realizar la evaluación son seleccionados los documentos que son considerados relevantes con respecto a los criterios que se proporcionaron en la consulta, tal proceso es desarrollado de acuerdo al Modelo de Recuperación de Información utilizado [12] [13].

Dado que los Modelos de Recuperación de Información realizan comparación entre elementos de texto (palabras) de documentos y consultas para determinar la relevancia de los documentos solicitados, estos modelos pueden ser aplicados para comparar Redes Semánticas Naturales, ya que se trata también de comparación de definidores expresados en palabras.

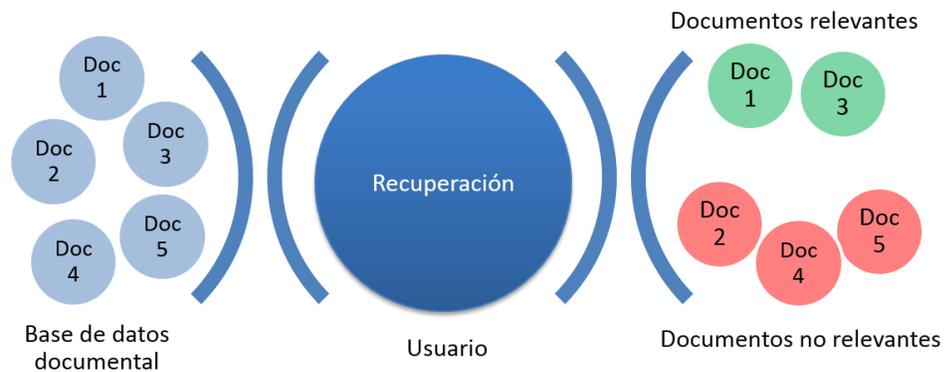


Figura 1. 2. Proceso de Recuperación de Información.

1.2. Justificación

Actualmente las RSN son usadas ampliamente en diferentes áreas de investigación y aplicación debido a su alcance en la representación del conocimiento, en donde una red acerca de una temática puede reflejar la representación de una muestra; es de interés para los investigadores comparar redes de diferentes poblaciones acerca de una misma temática, actualmente se utiliza un indicador llamado valor Q que representa el consenso grupal, es un índice de semejanza calculado con base en la coincidencia de los definidores principales entre las redes tomando en cuenta su posicionamiento [3].

Para mejorar el alcance y optimizar la forma de comparar dos redes es necesario realizar un estudio que permita diseñar un nuevo modelo de comparaciones desde una perspectiva de recuperación de información, que tenga como base el cálculo de similitud entre las redes; asimismo es importante generar un método de visualización de las RSN que permita representar a las redes así como a la similitud entre ellas.

1.3. Definición del Problema

Los trabajos acerca de RSN se han dirigido principalmente para aplicaciones de psicología social con el objetivo de investigar acerca de los significados, diversos estudios muestran que esta técnica ha sido utilizada para representar significados con distintos temas [14] [15] [16] [17].

Sin embargo existen también investigaciones de RSN en el campo de la evaluación cognitiva en las que se comparan las redes generadas, ya sea entre participantes del mismo grupo, o de participantes de diferentes grupos [18] , de participantes con expertos [19] [20] [21] o incluso del mismo participante antes y después de un curso [22]; para las investigaciones antes mencionadas se requiere hacer comparaciones entre las redes semánticas generadas, los métodos aplicados en esos trabajos para dicha comparación son: el cálculo del valor Q y el conteo de conceptos principales y relacionados que coinciden entre las redes.

En algunos trabajos se menciona que el cálculo del valor Q no representa en algunos casos un resultado representativo en la comparación de las redes [19] [23] y se cuestiona si la manera de calcular el indicador de comparación entre redes semánticas naturales es apropiado [24] [25].

Los sistemas de recuperación de información (*information retrieval systems*) procesan archivos de registros y solicitudes de información, identifican y recuperan de los archivos ciertos registros como respuesta a las solicitudes de información; la recuperación de ciertos registros en particular depende de la similitud entre los registros y las consultas definidas por la solicitud de información [26].

Hay tres procesos básicos que deben soportar los sistemas de recuperación de información: la representación del contenido de los documentos, la representación de las necesidades de información a buscar y la comparación entre las dos representaciones; la representación de la

información solicitada se denomina *query*; la comparación del *query* contra las representaciones de los documentos es llamado proceso de pareo (*matching*) [27]. Estos sistemas, aplican diversos modelos basados en similitud para comparar lo que se desea ser recuperado con los diversos documentos en donde se buscará con la finalidad de determinar cuál de ellos puede cumplir con la solicitud, estos modelos tienen posibilidad de ser aplicados para comparar redes semánticas naturales.

1.4. Objetivos

1.4.1. Objetivo General

Aplicar y evaluar un modelo de recuperación de información para realizar la comparación entre Redes Semánticas Naturales así como una representación visual de las mismas.

1.4.2. Objetivos Particulares

- Aplicar la técnica de Redes de Semánticas Naturales con una temática de ecología a estudiantes y expertos para generar redes a utilizar en este estudio.
- Generar representación visual de las RSN creadas.
- Seleccionar el Modelo de Recuperación de Información que se utilizará para comparar las RSN.
- Procesar la información de las RSN para aplicarla al Modelo de Recuperación de Información definido.
- Comparar las redes semánticas naturales utilizando el Modelo de Recuperación de Información seleccionado.
- Evaluar los resultados de la comparación entre redes semánticas naturales.
- Analizar las distintas medidas de similitud calculadas.

1.5. Hipótesis

Se ha definido una hipótesis general y un conjunto de hipótesis derivadas que contribuyen a la demostración de la hipótesis general. La hipótesis general a comprobar en esta tesis es:

- Es posible comparar dos redes semánticas naturales mediante un Modelo de Recuperación de Información.

De la hipótesis anterior se derivan las siguientes hipótesis las cuáles también son demostradas en este trabajo:

- Es posible comparar RSN de diferente cantidad de participantes mediante un Modelo de Recuperación de Información.
- Se puede adaptar un Modelo de Recuperación de Información para comparar RSN.
- Es posible identificar la medida de similitud adecuada para la comparación de RSN mediante un Modelo de Recuperación de Información.
- Se pueden representar visualmente las RSN mediante nubes de etiquetas diferenciando los definidores de acuerdo a distintos rangos de valores.
- Es posible comparar dos RSN de manera cualitativa mediante una representación visual con nube de etiquetas.

1.6. Preguntas de investigación

- ¿Es posible comparar dos redes semánticas naturales mediante un Modelo de Recuperación de Información?
- ¿Se pueden comparar dos RSN de diferente cantidad de participantes mediante un Modelo de Recuperación de Información?

- ¿Cuál es la medida de similitud adecuada para comparar RSN?
- ¿Se puede realizar una representación visual de RSN mediante nube de etiquetas diferenciando mediante colores distintos rangos de valores?
- ¿Se puede comparar de manera cualitativa dos RSN mediante su representación con nubes de etiquetas?

1.7. Estructura de la tesis

Esta tesis presenta la siguiente distribución: En el Capítulo 2 se expone el Marco Teórico de este trabajo, en donde se pueden encontrar los conceptos, modelos, teorías y trabajos relacionados en los que se sustenta esta tesis; en el Capítulo 3 es definida la Metodología que se siguió para comprobar la hipótesis propuesta; en el Capítulo 4 se muestran los resultados obtenidos, tanto en tablas como en representaciones visuales; las conclusiones de este estudio así como el posible trabajo futuro se exponen en el Capítulo 5.

CAPÍTULO 2

MARCO TEÓRICO

2.1. Representación del conocimiento

Existe una relación entre la investigación en Inteligencia Artificial (IA) con la filosofía desde sus inicios, ya que ambas se refieren a los problemas esenciales: ¿Qué es el significado? ¿Cómo lo sabemos?; estos problemas también son las principales preocupaciones del estudio de la semántica [28].

La principal diferencia entre los humanos y los sistemas inteligentes es que los seres humanos tienen una comprensión semántica y las máquinas no la tienen; ayudar a que las máquinas establezcan una visión del mundo es esencial para comprender IA a nivel humano; una comprensión semántica ayuda a las máquinas a establecer imágenes semánticas básicas, la imagen ordinaria refleja estáticamente un objeto o escena, mientras que las imágenes semánticas reflejan dinámicamente clasificación, objeto, persona, relación, reglas e interacción [28].

En el área de IA se desarrollan herramientas conceptuales para la comprensión de la manera en que se procesa la información, y estas herramientas son indispensables a la psicología para entender de qué forma la información es procesada por los seres humanos, algunos sistemas particulares de IA se transforman en las herramientas de análisis y síntesis, así como de

apoyo para el progreso de la psicología en su propia misión personal que es entender la mente del hombre [29] .

La Representación del Conocimiento es un área de IA encargada del estudio de la manera en que el conocimiento puede ser representado simbólicamente y manipulado automáticamente por los programas de razonamiento; existen diferentes formas de estudiar esta área, una de ellas se basa en términos de un lenguaje de representación como el de la lógica simbólica (IA basado en la lógica), otra se fundamenta en términos de la especificación y desarrollo de grandes sistemas basados en el conocimiento (ingeniería del conocimiento y ontologías); un enfoque diferente se refiere a la representación del conocimiento en un entorno de ciencia cognitiva, en donde la atención se centra en los modelos de los estados mentales humanos [30].

La representación del conocimiento permite visualizar el conocimiento que se tiene con respecto a diferentes temáticas a través de técnicas de procesamiento de información para el estudio de patrones, estructura y acomodo de información [30].

2.1.1. Redes Semánticas

El concepto de una red semántica es bastante antiguo en la literatura de la ciencia cognitiva y la inteligencia artificial, este concepto ha sido desarrollado de diferentes maneras. La expresión red semántica como se usa actualmente, puede considerarse como el nombre de una familia de sistemas de representación en lugar de un formalismo único [31].

El término red semántica se remonta a la tesis de Ross Quillian realizada en 1968, en la que se introdujo por primera vez, como una manera de hablar de la organización de la memoria semántica humana, o de la memoria para los conceptos; en concreto, las redes semánticas se concibieron, como un formato de representación que permiten que los significados de las palabras

sean almacenados, por lo que el parecido de estos significados a los usados por los humanos es posible [32].

La red semántica más antigua que se conoce se elaboró en el siglo tercero por el filósofo griego Porfirio, quien la utiliza para ilustrar el método de definición de categorías de Aristóteles, en donde se especifica el tipo género o general y la diferenciación que distingue diferentes subtipos del supertipo mismo; la Figura 2.1 muestra una versión del árbol de Porfirio, que fue elaborado por el lógico Pedro de España (1329), ilustra las categorías bajo la sustancia, que se llama el género supremo o la categoría más general.

Los métodos de definición y de razonamiento de Aristóteles se siguen utilizando en temas como inteligencia artificial, lenguajes de programación orientados a objetos y todos los diccionarios desde los primeros tiempos hasta la actualidad; la primera llamada red semántica se desarrolló en 1961, en la Universidad de Cambridge, se creó una lista de 100 tipos de conceptos primitivos, como Popular, Cosa, Hacer y Ser [33].

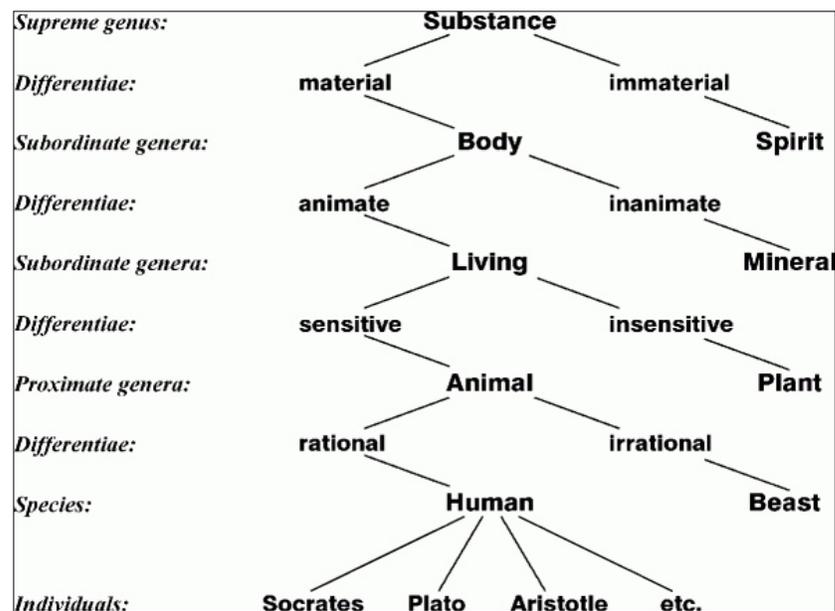


Figura 2. 1. Árbol de Porfirio, dibujado por Peter de España (1329) [34] .

En el desarrollo de las ciencias computacionales, la investigación en semántica ha tenido contribución en diversas áreas; la semántica formal estudia la especificación, operación y verificación de programas de computadora mediante herramientas matemáticas [35] [36]. Los modelos de datos relacionales se crearon para regular las organización y el comportamiento de sistemas computacionales [37]. La semántica informal trata de la modelización del mundo real mediante lenguajes como el UML (*Unified Modeling Language* Lenguaje Unificado de Modelado) [38] [39] y el ER (*Entity-Relationship* Entidad-Relación) [40]. Los métodos orientados a objetos intentan transformar modelos del mundo real en los programas mediante abstracciones semánticas básicas tales como clase, objeto, instancia, herencia, método, mensaje, encapsulamiento y polimorfismo [41] [42]. Los meta-lenguajes se utilizan para definir la semántica de lenguajes [43].

De acuerdo al enfoque de las redes semánticas, la generación del significado está dado por las relaciones entre los conceptos así como por la activación de los conceptos relacionados entre sí, y no por la esencia pura de los conceptos; una red semántica es un modelo de organización de información que permite estudiar la formación del significado aun y cuando no permita analizar la esencia pura de los conceptos que se involucran, lo anterior es importante, ya que las representaciones semánticas son las causantes de diversas formas de nuestra conducta [44].

2.2. Redes Semánticas Naturales

Las RSN están basadas en el campo de la investigación en aspectos de la memoria humana, establecen una visión cognitiva de la manera en que las personas representan la información así como del desarrollo de sus experiencias; la finalidad de esta técnica es estudiar la manera en que se establecen los significados desde la memoria semántica, una de sus características principales es que las propias redes determinan el significado de los conceptos [45].

Un principio fundamental de las redes semánticas es la existencia de una estructuración interna de la información que se encuentra en la memoria a largo plazo, en forma de red, en donde las palabras o eventos forman relaciones, las cuales en conjunto proporcionan el significado [46].

En otros métodos para construcción de redes semánticas, las definiciones conceptuales de la redes semántica son particulares debido a que son proporcionadas y definidas por el propio investigador, están dadas por categorías artificiales de conocimiento, lo anterior limita la obtención de un significado; mediante RSN se puede medir la distancia semántica, determinar la similitud semántica y permiten analizar procesos de inferencia semántica y trabajar con imágenes y sonido además de conceptos, por lo que esta técnica es útil para analizar la representación del conocimiento aprendido en un curso; actualmente existe potencial en la aplicación de la técnica de RSN en el uso de ontologías de conocimiento, en el área conocida como *Semantic Web* (Red de Internet Semántica) [44].

La técnica de RSN ha sido utilizada en diversas áreas desde su creación en 1981 en que Figueroa [3] plantea la creación de RSN para hacer investigación con humanos, esta técnica se ha actualizado y refinado hasta llegar a las categorías semánticas que incluyen una modificación a la técnica original de Figueroa, para analizar la forma en que se organizan significados a partir de la memoria semántica, mediante las palabras definidoras que utilizan los sujetos para definir una palabra estímulo, además consiguen definir numéricamente los atributos de la red y así realizar comparaciones [47].

Se describe de manera breve la técnica RSN, la cual se divide en dos partes fundamentales [48]:

- a) Solicitar a los sujetos participantes que definan con claridad y precisión a la palabra estímulo que se les presenta mediante el uso de palabras individuales que consideren que están relacionadas con ésta, a las palabras mencionadas se les llamará definidores; el tiempo que se

dará a los sujetos para definir cada concepto es de 90 segundos, debido a que si se les proporciona un tiempo mayor, los participantes tienden a proporcionar asociaciones libres en lugar de definiciones.

- b) Pedir a los participantes que jerarquicen, asignando el número 10 a la palabra más cercana o relacionada con la palabra estímulo, el número 9 a la que sigue en importancia, etc., se asignará 1 al concepto menos relevante, se deben jerarquizar todas las palabras dadas como definidores.

En la Tabla 2.1 se muestra un ejemplo de los definidores proporcionados por un participante al estímulo familia así como la calificación o valor semántico que le asignó a cada uno de ellos.

Tabla 2. 1. Ejemplo de definidores jerarquizados para el concepto estímulo familia mediante la técnica RSN.

Estímulo "familia"	
Definidores	Calificación
Apoyo	10
Unión	8
Fundamental	9
Hijos	2
Amor	6
Compartir	4
Hogar	5
Importante	7
Padres	3
Hermanos	1

En el apartado de 3.3 de este trabajo se describen los indicadores que se calculan una vez que se recopilan los definidores y su calificación.

2.3. Recuperación de Información

La necesidad de almacenar y recuperar información escrita se ha ido incrementando de manera importante durante siglos, especialmente después de la invención del papel y la imprenta; una vez que las computadoras se inventaron, la gente se percató de que podría utilizarlas para almacenar y recuperar grandes cantidades de información [49]. En 1945 Vannevar Bush publicó un artículo innovador en donde nace la idea de automatizar el acceso a gran cantidad de conocimiento almacenado [50]. En los años cincuenta esta idea se materializó en una descripción más concreta de la manera de automatizar la búsqueda en los archivos de texto, uno de los métodos más conocidos fue descrito por Luhn en 1957, en el cual propuso usar palabras como unidades de indexación para documentos y medir la coincidencia de palabras como un criterio para la recuperación [51]. Posteriormente varios desarrollos fundamentales se realizaron en la década de los sesenta, los más notables fueron realizados por Gerard Salton [52] y Cyril Cleverdon [53].

De acuerdo al sistema de clasificación del campo de estudio computacional definido por *Association for Computing Machinery* (ACM), el área de Recuperación de Información es clasificada en la categoría de Sistemas de Información y a su vez está subdividida en diversos tópicos que son mostrados en la Figura 2.2 [54].

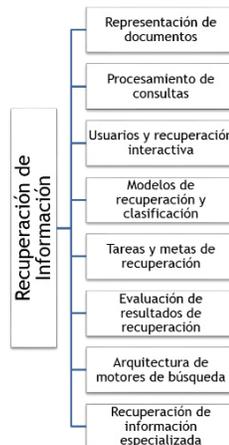


Figura 2. 2. Tópicos del área de estudio de Recuperación de Información.

2.3.1. Sistemas de Recuperación de Información

En los Sistemas de Recuperación de Información (SRI) la información que se procesa proviene de documentos; un SRI puede ser expresado como un conjunto de registros de información (documentos), un conjunto de peticiones (solicitudes) y algún proceso (similitud) que determine cuales registros cumplen con las necesidades de información expresadas por una solicitud o requerimiento [26]. En la Figura 2.3 se pueden observar las operaciones que se realizan en un sistema de recuperación de información.

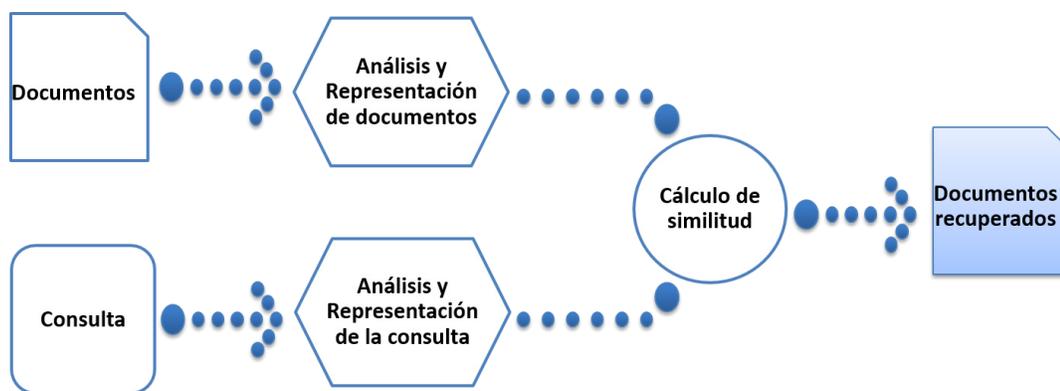


Figura 2. 3. Operaciones realizadas en un Sistema de Recuperación de Información.

Un SRI no obtiene una respuesta de una solicitud directamente, sino que localiza referencias a aquellos documentos que pueden contener información relacionada a la solicitud, actualmente existe un panorama más amplio, mediante el desarrollo de modelos para mejorar la funcionalidad de los SRI, tales como [55] :

- Modelos de Recuperación: El proceso de recuperación de información se puede modelar desde distintas perspectivas como estadística, el álgebra de Boole, el álgebra de vectores, la lógica difusa y el procesamiento del lenguaje natural.
- Clasificación: Proceso que etiqueta automáticamente los documentos con base en categorías definidas previamente.

- Agrupamiento: Procedimiento similar a la clasificación pero sin categorías definidas, el propio proceso determina las particiones.
- Extracción de Información: Proceso de extracción de fragmentos de texto que contengan una alta carga semántica y relaciona entre los fragmentos extraídos.
- Filtrado y Ruteo: Esta área apoya para la definición de perfiles de solicitudes de información de los usuarios además de analizar los nuevos documentos para dirigirlos a quienes estima que serán relevantes.
- Modelado de Usuarios: Mediante la interacción con los usuarios apoya para generar automáticamente perfiles que correspondan a sus necesidades de información.
- Recuperación de Información Multimedia: La recuperación de información se realiza con elementos diferentes al texto tales como imágenes o video.
- Recuperación *cross-language*: Localización de documentos escritos en cualquier lenguaje que son relevantes a una solicitud referida en otro lenguaje.
- Búsquedas Web: Operan su búsqueda sobre una red privada o pública, tiene características particulares tales como su tamaño y su dinamismo.

2.3.2. Modelos de Recuperación de Información

Al diseñar un SRI, se debe establecer de qué manera se realizará la representación de documentos así como de las solicitudes de información, y también debe determinarse cómo se van a comparar ambas representaciones, es decir, definir el modelo de recuperación sobre el que se desarrollará el sistema.

Los modelos para la recuperación de la información se clasifican de la siguiente manera [56]:

- Modelos clásicos. Son los modelos más utilizados: booleano, espacio vectorial y probabilístico.
- Modelos alternativos. Tienen como base la lógica difusa.
- Modelos lógicos basados en la lógica formal. Un proceso inferencial es utilizado para la recuperación de información.
- Modelos basados en la interactividad. Es posible expandir el alcance de la búsqueda, utilizan la retroalimentación de la relevancia de los documentos recuperados.
- Modelos basados en la Inteligencia Artificial. Están basados en redes neuronales, algoritmos genéticos y procesamiento del lenguaje natural.

2.3.2.1. Modelo Booleano

Es un modelo basado en teoría de conjuntos y álgebra booleana, utiliza expresiones booleanas (AND, OR, NOT) para la creación de consultas mediante las que establece el conjunto de condiciones que un documento debe satisfacer para considerarse relevante; es fácil de entender para un usuario de un SRI y se ha utilizado con frecuencia en sistemas comerciales, las consultas se materializan en expresiones booleanas en donde cada término de la consulta se identifica con el conjunto de los documentos que contienen dicho término, después se crean las intersecciones de conjuntos y se eligen aquellas que cumplen con las condiciones requeridas [57].

En un Modelo Booleano cada documento se representa por un conjunto de palabras clave, es asociado un peso binario a cada término del índice, se representa un 0 si el término está ausente en el documento y un 1 si el término aparece en el documento; las búsquedas consisten en expresiones formadas por palabras clave conectadas mediante operadores lógicos, el grado de similitud entre un documento y una consulta será binario, un grado de similitud 1

indica que el documento es relevante para la consulta y un grado de similitud 0 indica que no tiene relevancia [13].

Las desventajas del Modelo Booleano son las siguientes [58]:

- La recuperación basada en criterios de decisión binarios no permite una relación parcial entre la consulta y el documento.
- No se proporciona una calificación.
- La información necesita traducirse a una expresión booleana.
- El modelo con frecuencia retorna muy pocos o demasiados documentos en respuesta a una consulta del usuario.

Modelo Booleano Extendido.

El Modelo Booleano Extendido pretende aminorar algunas restricciones del Modelo Booleano Clásico, principalmente la capacidad de utilizar algún tipo de ponderación de términos y generar calificaciones de los documentos relevantes, está basado en la utilización de operadores lógicos en conjunto con modelos de representación vectoriales de los documentos, se utilizan técnicas de indización automática para ponderar los términos de los documentos, lo cual se suele hacer entre 0 y 1, un término representativo para un documento obtendría un valor cercano a 1, y uno con una baja representatividad obtendría un valor cercano a 0 o incluso 0 si no estuviera presente en el documento [9].

2.3.2.2. Modelo del Espacio Vectorial

El Modelo del Espacio Vectorial (MEV) es un modelo muy utilizado por los SRI, una de sus principales características es el uso de técnicas de ponderación de los términos en los documentos, con lo anterior supera la mayor debilidad del Modelo Booleano ya que posibilita ponderar la importancia de un término en un documento específico y permitir la coincidencia parcial entre el documento y la consulta en el proceso de recuperación, también facilita la calificación de la relevancia de los documentos con respecto de las consultas [26].

En este modelo el texto es representado por un vector de términos, los términos comúnmente son palabras; cualquier texto puede ser representado por un vector en un espacio dimensional [59]. En el MEV los documentos se representan a partir de vectores, de la siguiente manera [26]:

$$\vec{d}_j = (w_{1j}, w_{2j}, w_{3j}, w_{4j}, \dots, w_{nj}) \quad (2.1)$$

Donde n es igual al número total de elementos de representación considerados y por su parte w indica el peso que el término en concreto tiene para el documento j , el peso de un término es una medida de su importancia en la representación del documento.

Esta representación puede compararse de manera más precisa que las representaciones binarias; las comparaciones se realizan de acuerdo a lo siguiente: dos vectores pueden ser representados en el hiperplano y pueden ser medidas las diferencias de dirección entre ambos de la misma forma que se haría en un plano bidimensional mediante la comparación del coseno del ángulo que forman [49]. En la Figura 2.4 se muestra la representación del MEV para la comparación entre la consulta y los documentos.

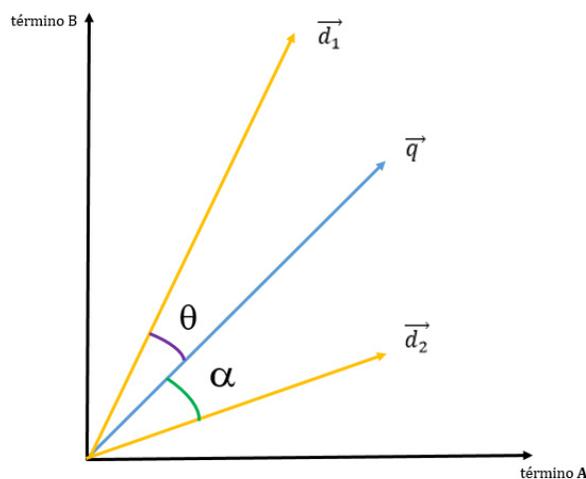


Figura 2. 4. El Modelo de Espacio Vectorial representa documentos y consultas como vectores para compararlos.

Para asignar una puntuación numérica a un documento para una consulta, este modelo mide la similitud entre el vector de consulta representado como \vec{q} y el vector del documento representado como \vec{d} , típicamente el ángulo entre dos vectores es usado como una medida de divergencia entre los vectores, y el coseno del ángulo es usado como la similitud numérica; como una alternativa, el producto punto entre dos vectores es usado para medir similitud, si todos los vectores son unitarios, entonces el coseno del ángulo entre dos vectores es el mismo que el producto punto [59].

La ecuación para medir el coseno de manera general (independientemente de la longitud y dimensiones de los vectores) se expresa como [26]:

$$\cos(\vec{a}, \vec{b}) = \frac{\vec{a} * \vec{b}}{|\vec{a}| * |\vec{b}|} \quad (2.2)$$

Existen diferentes funciones para medir la similitud entre documentos y consultas, las cuales están basadas en la consideración de ambos puntos en un espacio n -dimensional, estas medidas son el coeficiente de Dice, el coeficiente de Jaccard y la distancia euclídea, en el apartado 2.4 se detallan estas medidas de comparación.

2.3.2.3. Modelo Probabilístico

El Modelo Probabilístico utiliza la teoría de la probabilidad para construir la función de búsqueda, para definir como se aplicará y para estimar los valores de los parámetros asociados a esta función; la información para integrar la función de búsqueda se obtiene del conocimiento de la distribución de los términos de indización del conjunto de documentos [60]. La estimación es la parte clave del modelo, y aquí es donde la mayoría de los modelos probabilísticos difieren entre sí; se han propuesto varios modelos probabilísticos, cada uno basado en una técnica de estimación de probabilidad diferente [49].

Este modelo estima la probabilidad de que, dada una consulta q , un documento d sea relevante para esa consulta representado de la siguiente manera: $P(Rel | d)$, en el modelo se intenta obtener un conjunto de documentos relevantes (denominado R) que deberá maximizar la probabilidad de relevancia; un documento se considera relevante si su probabilidad de ser relevante, $P(Rel | d)$, es mayor que la probabilidad de no ser relevante, $P(noRel | d)$; para calcular la similitud de un documento con una consulta, $sim(q, d)$, se calcula la división entre ambas probabilidades [60].

2.4. Similitud

El conocimiento de similitud y disimilitud es necesario para los campos de minería de datos, reconocimiento de patrones e inteligencia artificial; sin embargo, las aplicaciones de estos conceptos no están limitadas al campo de ciencias computacionales, otros campos de ciencias naturales y sociales así como de estadística e ingeniería los utilizan [61].

El siglo veinte fue testigo de enormes esfuerzos para explotar las nuevas medidas de distancia/similitud para una variedad de aplicaciones, existe un número importante de estas medidas encontradas en diferentes campos como la antropología, la biología, la química, la informática, la ecología, la teoría de la información, la geología, las matemáticas, la física, la psicología y estadística [62].

La similitud es una medida numérica de qué tan parecidos son dos objetos de datos, es mayor cuando los objetos son más parecidos, generalmente se encuentra en el rango $[0,1]$; por otra parte la disimilitud es la medida numérica de lo diferente que son dos objetos de datos y disminuye cuando los objetos son más parecidos, el valor de disimilitud mínimo suele ser 0 y el superior varía; hay considerables esfuerzos en la búsqueda de las medidas apropiadas entre varias opciones, ya que es de fundamental

importancia para la clasificación de patrones, el agrupamiento, y los problemas de recuperación de información [63].

2.4.1. Medidas de Similitud

Para medir lo similar (o disimilar) que son objetos o individuos, existe una gran cantidad de índices de similitud y de disimilitud o divergencia, todos ellos tienen propiedades y utilidades distintas y habrá que ser consciente de ellas para su correcta aplicación de acuerdo al caso de estudio [64].

Si se tiene un conjunto de n objetos caracterizados por los valores de sus variables (x_1, x_2, \dots, x_n) , serán más similares cuanto más parecidos sean sus valores en las variables individuales, lo que se traducirá en que se sitúen próximos en el espacio n -dimensional, por lo tanto, cualquier método de agrupación comienza por el cálculo de la similitud entre los objetos; los tipos de medidas de similitud más importantes son: distancias y coeficientes de asociación y correlación [65].

2.4.1.1. Similitud Cosenoidal

Una medida de similitud está dada por el coseno del ángulo entre dos vectores, el modelo propone evaluar el grado de similitud entre un vector a y un vector b , como la correlación entre los dos, esta correlación puede medirse con el coseno del ángulo entre dichos vectores, lo que es equivalente al producto escalar normalizado entre ambos vectores como muestra la ecuación 2.3 [66].

$$SimCos(a,b) = \frac{\sum_{i=1}^k a_i * b_i}{\sqrt{\sum_{i=1}^k a_i^2} * \sqrt{\sum_{i=1}^k b_i^2}} \quad (2.3)$$

Esta ecuación produce un valor entre 0 y 1, en donde un valor de 0 indica la perpendicularidad (ortogonalidad) de los vectores, lo cual indica que se tendrían dos vectores en dirección completamente diferente, y un valor de 1 indica que los vectores tienen una dirección idéntica [9].

2.4.1.2. Coeficiente de Jaccard

Está basado en el coeficiente de similitud del coseno, su aplicación centrada en usos estadísticos, también se aplica a recuperación de información y mide la similitud entre conjuntos; se puede definir como el tamaño de la intersección (numerador) dividido por el tamaño de la unión de la muestra, en este caso la suma de los pesos al cuadrado del documento d y la consulta q menos la intersección [67].

$$SimJacc(d_{(d)}, q) = \frac{\sum_{n=1} (P_{(n,d)} * P_{(n,q)})}{\sum_{n=1} (P_{(n,d)})^2 + \sum_{n=1} (P_{(n,q)})^2 - \sum_{n=1} (P_{(n,d)} * P_{(n,q)})} \quad (2.4)$$

2.4.1.3. Coeficiente de Dice

Es una adaptación del cálculo del coeficiente del coseno, la diferencia está en que la cardinalidad del numerador es 2 veces la información compartida y el denominador es la suma de los pesos al cuadrado del documento y la consulta; para los conjuntos de palabras clave utilizadas en la recuperación de información, el coeficiente puede ser definido como dos veces la información compartida (intersección) sobre la suma de cardinalidades [68].

$$SimDice(d_{(d)}, q) = \left(\frac{2 \sum_{n=1} (P_{(n,d)} * P_{(n,q)})}{\sum_{n=1} (P_{(n,d)})^2 + \sum_{n=1} (P_{(n,q)})^2} \right) \quad (2.5)$$

2.4.2. Distancias como indicadores de disimilitud

Distancia o disimilitud entre dos individuos i y j es una medida indicada por $d(i,j)$, que mide el grado de desemejanza entre ambos objetos o individuos con relación a un cierto número de características cuantitativas y/o cualitativas; el valor de $d(i,j)$ es siempre un valor no negativo, y cuanto mayor sea este valor mayor será la diferencia entre los individuos i y j , toda distancia debe verificar, al menos, las siguientes propiedades [63]:

$$d(i,j) > 0 \text{ (no negatividad)}$$

$$d(i,i) = 0$$

$$d(i,j) = d(j,i) \text{ (simetría)}$$

Existe una gran cantidad de distancias e indicadores de disimilitud y no se puede disponer de una regla general que permita definirla de manera conveniente para todo tipo de análisis; de las propiedades que posea, de la naturaleza de las variables utilizadas, de los individuos estudiados y de la finalidad del análisis dependerá la adecuada elección de una u otra [69].

2.4.3. Familia Minkowski de disimilitud

Hace dos mil años, Euclides afirmó que la distancia más corta entre dos puntos es una línea y por lo tanto la ecuación (1) en la Tabla 2.2 es conocida predominantemente como la distancia euclidiana, a menudo se llama métrica de Pitágoras ya que se deriva del teorema de Pitágoras; en el siglo diecinueve, Hermann Minkowski considera la distancia cuadra de ciudad (*City Block*), otros nombres para la ecuación (2) incluyen la distancia rectilínea, norma de taxi (*taxicab norm*), y la distancia de Manhattan; Hermann también generaliza las fórmulas (1) y (2) a la ecuación (3), que está acuñada después de Minkowski, cuando p va al infinito, la ecuación (4) puede ser derivada y se llama la distancia de tablero de ajedrez en 2D, la aproximación minimax, o la distancia de Chebyshev [69].

Tabla 2. 2. Familia Minkowski de medidas de disimilitud.

1. Euclideana L_2	$d_{Euc} = \sqrt{\sum_{i=1}^d P_i - Q_i ^2}$	1)
2. City block L_1	$d_{CB} = \sum_{i=1}^d P_i - Q_i $	2)
3. Minkowski L_p	$d_{Mk} = \sqrt[p]{\sum_{i=1}^d P_i - Q_i ^p}$	3)
4. Chebyshev L_∞	$d_{Cheb} = \max_i P_i - Q_i $	4)

2.5. Visualización de Redes Semánticas Naturales

La visualización de la información actualmente se define como el uso de apoyo computacional para representar datos con el objetivo de incrementar la cognición; es aplicada para facilitar la comprensión del conocimiento [70].

Se han utilizado diversos métodos para realizar la representación visual de las RSN, en el trabajo realizado por Zermeño, Arellano y Ramírez [71] en el que representan el significado de televisión y tecnologías de la información, utilizan mapas conceptuales, definidos como la representación visual que permite que la expresión de conocimiento pueda ser fácilmente entendida por otros, los conceptos se encierran en círculos o rectángulos que se conectan o relacionan por líneas o arcos [72]. La Figura 2.5 muestra la RSN representada por un mapa conceptual.

Otras representaciones visuales son realizadas con esquemas conceptuales sencillos que relacionan conceptos, como en el trabajo realizado para representar los conceptos de presión y flotación en estudiantes de bachillerato [73] y en el estudio para representar las percepciones del clima escolar en una institución como lo muestra la Figura 2.6.

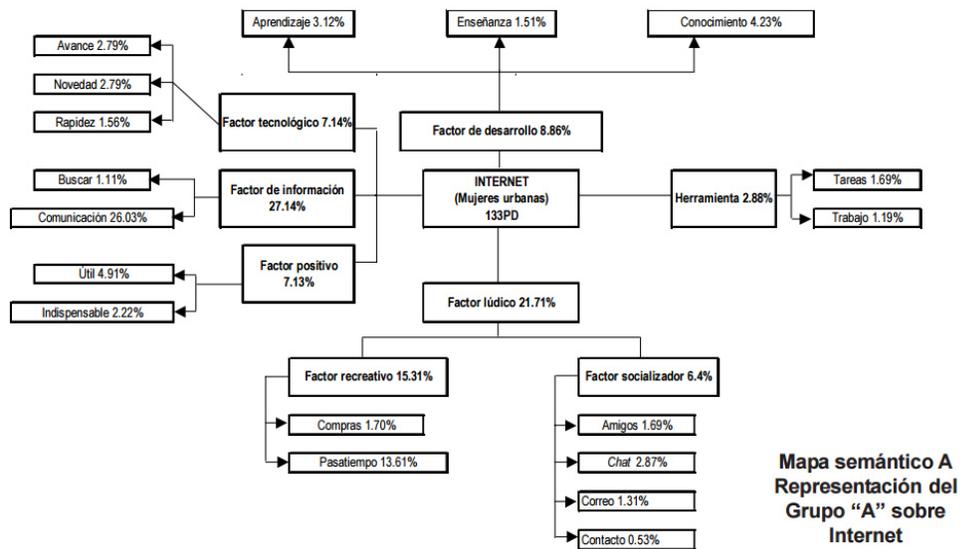


Figura 2.5. Visualización de RSN mediante mapa conceptual [71].

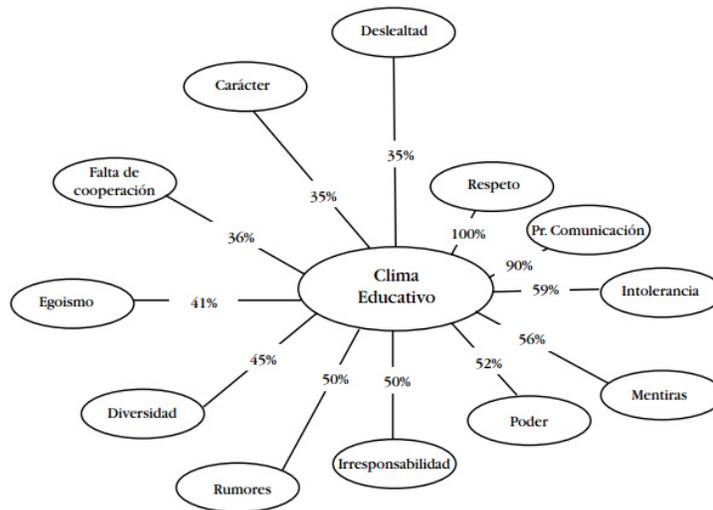


Figura 2.6. Visualización de la red semántica natural para representación del clima escolar [15].

En otros trabajos se aplica una técnica denominada Pathfinder que estima la proximidad entre elementos conectados [74]. Es posible hacer la representación de la red Pathfinder mediante una herramienta llamada KNOT

CAPÍTULO 3

METODOLOGÍA

3.1. Descripción del estudio

En este trabajo se generan Redes Semánticas Naturales para la propia investigación mediante una aplicación web conectada a una base de datos, además se desarrolló una definición matemática formal para los indicadores que son calculados en la técnica RSN.

De los Modelos de Recuperación de Información analizados se seleccionó el Modelo de Espacio Vectorial para realizar la comparación entre las RSN ya que éste considera el peso o importancia de los términos y puede proporcionar un valor de similitud relativo al parecido entre los documentos a comparar; para la comparación se realizó el cálculo de tres medidas de similitud: el coeficiente de similitud cosenoidal el coeficiente de Jaccard y el coeficiente de Dice.

Se construyeron visualizaciones de RSN mediante nubes de etiquetas diferenciadas por colores con el objetivo de percibir el contenido de la RSN así como de identificar su similitud.

A continuación se especifican las actividades que se realizaron para cumplir con el objetivo de la investigación, las cuales serán detalladas posteriormente:

- a. Aplicar la técnica de RSN con el objetivo de generar las redes semánticas a utilizar como material de comparación.
- b. Realizar el pre-procesamiento de las RSN generadas para la preparación para la comparación.
- c. Construir la visualización de las RSN mediante una nube de *tags* (etiquetas).
- d. Comparar las RSN mediante el Modelo del Espacio Vectorial.
- e. Análisis de resultados de comparación.

3.2. Participantes

Para la generación de las RSN participaron 163 individuos seleccionados intencionalmente o de conveniencia; 159 de estos participantes son estudiantes de ingeniería de entre 18 y 21 años, familiarizados con el uso de internet y 4 de los participantes son investigadores expertos en la temática seleccionada.

La temática seleccionada para la representación de significado mediante las RSN fue Ecología, estos estudiantes son considerados como principiantes en la temática mencionada ya que sólo tienen la formación básica relacionada con ecología, es decir conocimientos generales del tema. Cuatro participantes son personas especializadas en ecología con las que se representó la red de expertos en la temática. La Tabla 3.1 muestra la cantidad de participantes por RSN generada.

Tabla 3.1. Representa la cantidad de participantes por cada RSN generada.

RSN GENERADA	CANTIDAD DE PARTICIPANTES
ECOLOGÍA 1	32
ECOLOGÍA 2	47
ECOLOGÍA 3	35
ECOLOGÍA 4	45
EXPERTOS	4

3.3. Especificación de técnica para realizar estudio de Redes Semánticas Naturales.

Como se describió en el apartado 2.2, la técnica de RSN consiste en que los sujetos proporcionen definidores relacionados a una palabra estímulo, a los cuáles les asignan un valor o peso que depende de la cercanía del significado a esa palabra estímulo proporcionada. Enseguida se describen los indicadores definidos para esta técnica así como la representación formal matemática de sus indicadores.

Una vez que se tienen los definidores de la palabra estímulo para la cantidad de sujetos establecida que participa en el estudio, se genera una tabla de frecuencias representada como una matriz, en donde sus filas son los definidores mencionados para un concepto estímulo y sus columnas son los valores semánticos, esta matriz se denomina P ya que en ella se calcula el peso semántico tomando en cuenta la frecuencia con que fue mencionado cada uno de los definidores en cada valor semántico como lo muestra la Tabla 3.2., en donde se representa un ejemplo de los definidores y valores semánticos proporcionados por diez participantes a la palabra estímulo contaminación.

El paso siguiente consiste en calcular los indicadores propuestos por Figueroa, González y Solís [3] y Valdez [46] que se definen a continuación, se incluye una representación matemática formal para cada uno de ellos:

Valor J: Son todas las palabras (definidores) utilizadas por los sujetos para definir los conceptos que actuaron como estímulo.

Cada sujeto puede verse en términos del conjunto de sus definidores. Formalmente se tendría un conjunto $S_i = \{def_1, def_2, \dots, def_n\}$ donde S_i es el sujeto en cuestión, cada def_i representa uno de sus definidores y n es la cantidad de los mismos.

El conjunto de definidores D , entonces, está dado por la unión de los sujetos, tal que $D = S_1 \cup S_2, \dots, S_k$ donde k es la cantidad de sujetos. El valor J corresponde al tamaño de D , es decir $J = |D|$ e indica la riqueza a nivel semántico que tiene la red.

Tabla 3.2. Matriz P que almacena la frecuencia acumulada definidor/jerarquía y cálculo de Valor M para 10 participantes.

Jerarquía	1	2	3	4	5	6	7	8	9	10		
Valor semántico	10	9	8	7	6	5	4	3	2	1		
Definidor	Valor M											
basura	2*10	2*9	2*8	1*7							20+18+16+7	61
fábrica	3*10		2*8		2*6		1*4				30+16+12+4	52
gris		1*9		2*7		1*5		1*3			9+14+5+3	31
auto	2*10			2*7		1*5	1*4				20+14+5+4	43
ruido		2*9	3*8		1*6	1*5					18+24+6+5	53
enfermedad				2*7	2*6	1*5	1*4			1*1	14+12+5+4+1	36
reciclaje		2*9			3*6					1*1	18+18+1	37
smog	1*10	3*9	1*8			2*5					10+27+8+10	55
anuncio							2*4				8	8
tóxico					2*6			1*3	1*2		12+3+2	17
ozono	1*10						1*4				10+4	14
humo		1*9	1*8			2*5		1*3			9+8+10+3	30

Valor M: Se calcula tomando en cuenta el peso semántico asignado por los sujetos a cada concepto definidor de acuerdo a su importancia o pertinencia.

En la Matriz P de la Tabla 3.2, las filas representan todos los definidores mencionados por los sujetos, denotado por D , como se mencionó antes, el total de éstos está dado por J ; las columnas representan el valor semántico V , dichos valores están definidos del 10 al 1 y corresponden con la

jerarquía asignada por los sujetos a cada definidor del 1 al 10, esto es, al nivel de jerarquía 1 le corresponde un peso semántico de 10, a la jerarquía 2 le corresponde un peso semántico de 9, etc. Cada celda en la matriz almacena la frecuencia acumulada en la que fue mencionado un definidor j en la jerarquía i multiplicada por el valor semántico que les corresponde como lo expresa la siguiente fórmula para el cálculo del peso semántico P de cada celda:

$$P_{j,i} = f(def_{j,i}) * V_i \quad (3.1)$$

Donde

$$j = 1 \dots J$$

$$i = 1 \dots 10$$

$f(w_{j,i})$ Representa la frecuencia del concepto definidor j en la jerarquía i .

V_i Representa el valor semántico correspondiente a la jerarquía i .

El Valor M es una medida de la relevancia de cada concepto como un definidor del concepto meta y se calcula para cada definidor def mediante la fórmula 3.2.

$$M_j = \sum_{i=1}^{10} P_{j,i} \quad (3.2)$$

En la Tabla 3.2 se puede observar el cálculo del Valor M para cada definidor def .

Grupo SAM (*Semantic Association Memory*): Está formado por los 10 definidores def con el Valor M más elevado de cada uno de los conceptos que actúan como estímulo. Forman el significado de la red sobre el concepto estímulo. En la Tabla 3.3 se muestra el Grupo SAM correspondiente al ejemplo representado en la Tabla 3.2

Valor G: Se refiere a la cercanía de los valores M en cada uno de los grupos SAM. Cuando el valor G es elevado representa baja densidad semántica y cuando el valor G es pequeño muestra en el grupo una densidad semántica

alta, se calcula aplicando la fórmula 3.3, para este ejemplo el valor G es de 4.4, si este valor estuviera normalizado sería equivalente a 0.73, por lo que se considera que la densidad semántica es baja.

$$G = \frac{M_{Max} - M_{Min}}{10} \quad (3.3)$$

Donde

M_{Max} es el valor máximo calculado de M_j

M_{Min} es el valor mínimo calculado de M_j

Tabla 3. 3. Muestra la selección del Grupo SAM y el Valor FMG de cada definidor en el Grupo SAM.

Definidor	Valor M	Valor FMG
basura	61	100
smog	55	90.16
ruido	53	86.88
fábrica	52	85.24
auto	43	70.49
reciclaje	37	60.65
enfermedad	36	59.01
gris	31	50.81
humo	30	49.18
tóxico	17	27.86
ozono	14	
anuncio	8	

Grupo SAM

Valor FMG. Proporciona la distancia semántica que hay entre los definidores de un Grupo SAM, es decir, el porcentaje de ponderación correspondiente al valor M de cada definidor en el grupo SAM, con respecto al valor más alto de M en tal grupo.

Se obtiene en función del definidor con el puntaje M_{Max} , el cual siempre representará el 100%. Para calcular el valor FMG de cada definidor en cada Grupo SAM se toma el definidor que tiene el valor M más alto, si los valores M_j del Grupo SAM están ordenados de manera creciente, entonces $M_{Max} = M_1$, el cálculo se realiza aplicando la fórmula 3.4.

$$FMG_j = \frac{M_j * 100}{M_1} \quad (3.4)$$

Nótese que siempre $FMG_1=100$; el valor FMG permite calcular una distancia estandarizada entre el definidor con Valor M más alto con respecto a los otros. El valor G es el factor de estandarización de la diferencia. En la Tabla 3.3 se muestra el cálculo de FMG para cada definidor en el Grupo SAM.

3.4. Generación de Redes Semánticas Naturales

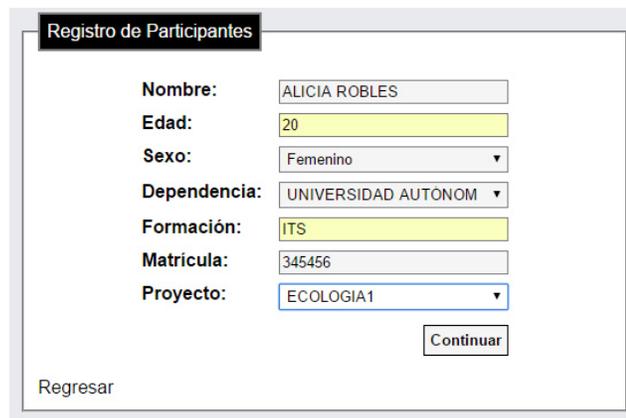
Los conceptos estímulo utilizados en este estudio para la generación de la RSN fueron: contaminación, reciclaje, ciudad, reforestación, ecología, medio ambiente y naturaleza. Como se definió en el apartado 2.2, en la técnica de Redes Semánticas Naturales, es necesario que los participantes escriban conceptos asociados (definidores) a los conceptos estímulo que se les presentan y que posteriormente les asignen una calificación a cada definidor.

Aunque existe una aplicación computacional que se utiliza para realizar este proceso de aplicación y generación de RSN llamada SEMNET [77], para realizar este proceso se utilizó una aplicación web denominada *RSNweb* [78] desarrollada por el grupo de investigación de ingeniería cognitiva de la FIME, a diferencia de SEMNET que es una aplicación de escritorio desarrollada en VisualBasic, *RSNweb* fue desarrollada para optimizar el proceso de aplicación de la técnica de RSN, ya que no es necesario instalarse en cada equipo ni recuperar la información manualmente de cada computadora en la que se realiza, sino que puede ser accedida desde cualquier equipo de cómputo que

cuenta con acceso a internet, fue desarrollada en lenguaje PHP, jquery y almacena la información en una base de datos MySQL.

3.4.1. Aplicación de RSN a principiantes utilizando *RSNweb*

El proceso de aplicación de la técnica de RSN se realizó en cuatro grupos de clase de la carrera ITS de la FIME, cada alumno cuenta con una computadora personal para realizar la prueba. Se solicitó al estudiante que accediera a un enlace de la aplicación que le fue proporcionado para iniciar con el registro de sus datos personales como lo muestra la Figura 3.1.



Registro de Participantes	
Nombre:	ALICIA ROBLES
Edad:	20
Sexo:	Femenino
Dependencia:	UNIVERSIDAD AUTONOM
Formación:	ITS
Matricula:	345456
Proyecto:	ECOLOGIA1
<input type="button" value="Continuar"/>	
Regresar	

Figura 3. 1. Muestra la pantalla de inicio de *RSNweb* para el registro de la información de los participantes en el estudio.

De acuerdo a la técnica de RSN, el siguiente paso es mostrar al participante una pantalla con las instrucciones a seguir, en donde se explica de qué manera debe realizar la prueba, esta pantalla también incluye un ejemplo de las acciones a realizar, como se expone en la Figura 3.2.

INSTRUCCIONES

La tarea a realizar es definir un concepto por medio de verbos, sustantivos y adjetivos. No se pueden utilizar artículos, ni frases completas.

Se van a presentar 7 conceptos asociados a:

ECOLOGIA1

Tendrá 60 segundos para escribir los definidores, por lo que se le pide que sea lo más rápido y preciso que pueda. Es importante que al finalizar cada palabra oprima la tecla enter. Una vez que termine el tiempo para escribir definidores, tendrá que calificar cada uno de estos en una escala de 1 a 10, donde 10 es la palabra que mejor define y 1 es la que menos lo define. Las calificaciones se pueden repetir tantas veces sea necesario.

Por ejemplo, si el concepto a definir es MANZANA, entonces usará otros conceptos como ARBOL, FRUTA, ROJA. Recuerde que NO podrá utilizar artículos(el, la, los, ...) para formar frases como "LA MANZANA ES ROJA", ni tampoco podrá utilizar frases completas como "CAEN DE LOS ARBOLES". No olvide que solo podrá utilizar conceptos como verbos(comer, caminar, ...), adjetivos(roja, amarilla, ...) o sustantivos(recolector, cocinero, ...).

0

MANZANA

ARBOL	10
FRUTA	9
ROJA	7
MARCA	7

De click en el botón para comenzar con el ejercicio de práctica.

Figura 3. 2. Pantalla de Instrucciones que muestra *RSNweb* al participante.

Una vez que son mostradas las instrucciones, se solicita al participante que realice una práctica con un concepto estímulo de ejemplo, que en este caso es la palabra “perro”, lo anterior es con el objetivo de que se familiarice con la técnica antes de iniciar con los estímulos correspondientes al estudio; después de realizar el ejercicio de práctica se empiezan presentar al participante los conceptos estímulo para que escriba palabras asociadas a estos, el participante dispondrá de 60 segundos para tal acción; posteriormente se le solicita que califique cada uno de los definidores de acuerdo a su cercanía o relación con el concepto estímulo, como lo muestra la Figura 3.3.

Cabe mencionar que las palabras estímulo aparecerán de manera aleatoria, es decir, no aparecerán en el mismo orden para todos los participantes. Una vez que se responde a los siete conceptos estímulo se agradece al participante y termina la prueba.

a) 9

MEDIO AMBIENTE

VERDE	
ECOLOGIA	
BOSQUE	
HABITAT	
FLORA	
FAUNA	

b) 0

MEDIO AMBIENTE

VERDE	9
ECOLOGIA	10
BOSQUE	9
HABITAT	10
FLORA	8
FAUNA	8

Figura 3. 3. Interfaz de la aplicación a) Muestra en la pantalla el concepto estímulo para que el participante escriba definidores asociados, en la esquina superior izquierda aparece el tiempo restante b) Permite calificar cada definidor de acuerdo a la cercanía con la palabra estímulo.

3.4.2. Limpieza de información y generación de RSN de principiantes.

Los definidores proporcionados son almacenados en una base de datos relacional, la cual es accedida por un programa que permite hacer la limpieza de los definidores, en este proceso se corrige la ortografía de los conceptos y se estandarizan de acuerdo a criterios establecidos tales como: los conceptos deben estar expresados en singular y en género masculino con el objetivo de que sean asociados cuando se contabilicen los definidores similares. La Figura 3.4 muestra la interfaz en la que se realiza el proceso de limpieza.

DEFINIDOR	OPCIONES
ATMOSFERA	<input type="checkbox"/>
AUTOBUS	<input type="checkbox"/>
AUTOPISTAS	<input checked="" type="checkbox"/> <input checked="" type="checkbox"/>
AVENIDA	<input type="checkbox"/>

Figura 3. 4. Interfaz de *RSNweb* para el proceso de limpieza de definidores de la red.

Una vez que se concluye el proceso de limpieza se realiza la identificación de los grupos SAM de la red, tomando los primeros diez definidores con el valor M más alto, se identifica también el valor J denotado para la cantidad de definidores diferentes por estímulo y por red.

La Figura 3.5 muestra el formato generado con la información recopilada para la red ECOLOGÍA 1, la representación incluye el grupo SAM de cada estímulo, el valor J y el valor G por grupo SAM, además incluye el valor M obtenido del cálculo de la sumatoria de la frecuencia de cada definidor multiplicada por su valor semántico de acuerdo a la fórmula 3.2.

3.5. Generación de RSN de expertos

Para obtener las redes de expertos se contactó a investigadores de la Facultad de Ciencias Biológicas de la Universidad Autónoma de Nuevo León que son especialistas en Ecología, en la Tabla 3.4 se puede observar el área de investigación desarrollada para cada uno de los expertos participantes

PROYECTO - ECOLOGIA 1

CIUDAD			CONTAMINACION			ECOLOGIA			MEDIO AMBIENTE		
F	Definidor	M	F	Definidor	M	F	Definidor	M	F	Definidor	M
1	EDIFICIO	218	2	BASURA	243	4	VERDE	119	2	NATURAL	133
2	CARRO	198	1	HUMO	109	4	ARBOL	117	4	ARBOL	117
2	CONTAMINACIÓN	142	1	SUCIO	101	1	RECICLAJE	109	3	ANIMAL	96
1	PERSONA	137	1	ENFERMEDAD	76	4	PLANTA	102	2	ECOSISTEMA	96
1	CASA	115	1	SMOG	71	2	NATURAL	92	4	PLANTA	92
1	TRAFICO	82	2	CARRO	66	4	AMBIENTE	86	3	ECOLOGÍA	70
1	METROPOLI	66	1	DESECHO	53	3	ANIMAL	79	2	CONTAMINACIÓN	57
1	GENTE	56	4	AMBIENTE	47	1	LIMPIA	77	4	VERDE	51
1	CALLE	49	1	FABRICA	42	2	ECOSISTEMA	53	2	AGUA	49
1	TIENDA	33	1	TOXICO	33	1	CUIDADO	51	2	AIRE	49
J = 97			J = 79			J = 66			J = 85		
G = 18.5			G = 21			G = 6.8			G = 8.4		

NATURALEZA			RECICLAJE			REFORESTACION		
F	Definidor	M	F	Definidor	M	F	Definidor	M
3	ANIMAL	199	1	REUSAR	202	4	ARBOL	270
4	PLANTA	150	2	BASURA	176	4	PLANTA	171
4	ARBOL	136	1	PLÁSTICO	125	2	BOSQUE	126
4	VERDE	115	1	PAPEL	100	1	OXÍGENO	86
2	AGUA	85	4	AMBIENTE	66	3	ECOLOGÍA	70
1	FLORA	75	1	LATA	55	4	AMBIENTE	59
2	BOSQUE	70	3	ECOLOGÍA	55	1	TIERRA	58
1	VIDA	63	1	ORGANICO	52	4	VERDE	47
1	FAUNA	56	1	BOTELLA	45	1	SEMILLA	45
1	RIO	55	1	BOTE	45	2	AIRE	35
J = 85			J = 59			J = 73		
G = 14.4			G = 15.7			G = 23.5		

Figura 3. 5. Grupos SAM de RSN del Proyecto ECOLOGÍA1, la columna F representa la frecuencia de aparición del definidor y la columna M representa el valor M calculado para cada definidor.

Tabla 3. 4. Muestra el área de especialidad de los expertos en Ecología que contribuyeron al estudio.

	Especialidad
Experto 1	Educación ambiental
Experto 2	Ecología forestal
Experto 3	Legislación ambiental
Experto 4	Ecología acuática

3.5.1. Aplicación de RSN a expertos

Para obtener la contribución individual de los expertos se le solicitó a cada uno de ellos que completara una plantilla con los definidores asociados a las palabras estímulo y que, como lo hicieron los principiantes, calificaran a cada definidor escrito de acuerdo a su cercanía con el concepto estímulo. La Figura 3.6 muestra la plantilla que corresponde al Experto 1, en la cual aparece

el valor F que es la frecuencia de aparición de los conceptos y el valor M, que es la calificación de cada definidor; con respecto al valor J que corresponde a la cantidad de definidores proporcionados en cada grupo SAM, en este caso sería tendría el valor diez.

PLANTILLA EXPERTO 1

CIUDAD			CONTAMINACION			ECOLOGIA			MEDIO AMBIENTE		
F	Definidor	M	F	Definidor	M	F	Definidor	M	F	Definidor	M
1	ESTRES	5	1	HUMO	7	1	ORGANISMO	9	2	INTERACCIÓN	10
2	GENTE	7	1	BASURA	5	2	INTERACCIÓN	7	3	VERDE	2
2	CASA	8	1	DESCOMPOSICIÓN	9	1	NATURALEZA	8	1	FLUJO	8
1	CEMENTO	6	1	DETERIORO	10	1	MAR	3	2	CICLO	7
1	METROPOLI	9	1	QUÍMICOS	8	2	TIERRA	4	2	BIODIVERSIDAD	6
1	POBLACIÓN	10	1	SMOG	6	2	BIODIVERSIDAD	10	1	RELACIONES	4
1	CARRO	4	1	CONTAGIO	2	1	VIENTO	2	2	GENTE	1
2	FORTALEZA	2	1	MEZCLA	3	2	CASA	1	2	RECURSOS	5
1	PANORÁMICO	3	1	INFECCIÓN	1	1	SISTEMA	5	1	HÁBITAT	3
1	SEÑALIZACIÓN	1	1	PUDRICIÓN	4	2	RECURSOS	6	2	UNIVERSO	9

NATURALEZA			RECICLAJE			REFORESTACION			ENERGIA		
F	Definidor	M	F	Definidor	M	F	Definidor	M	F	Definidor	M
1	VIDA	10	1	TRANSFORMAR	9	1	PLANTA	9	1	SOL	8
3	VERDE	7	1	PROCESAR	10	3	VERDE	8	1	POTENCIA	5
2	UNIVERSO	9	1	APROVECHAR	3	1	SIERRA	5	1	INTENSIDAD	4
2	TIERRA	8	1	RETOMAR	4	1	PODAS	3	1	IMPULSO	3
1	ESTRUCTURA	6	1	REHABILITAR	7	1	INJERTO	2	2	FORTALEZA	9
1	ESENCIA	2	1	ENCAUSAR	2	1	TRANQUILIDAD	1	1	FUERTE	1
1	IDENTIDAD	1	1	RECUPERAR	8	1	ARBOLAR	6	1	PODER	7
1	VARIEDAD	5	1	REUSAR	6	1	BOSQUE	7	1	LUZ	6
1	ARMONÍA	3	1	REDUCIR	5	1	HOJA	4	1	NUCLEO	2
2	CICLO	4	1	COMPOSTA	1	2	SUSTENTABILIDAD	10	2	SUSTENTABILIDAD	10

Figura 3. 6. Muestra la Plantilla que representa la RSN del Experto 1.

3.6. Modelo de Espacio Vectorial en Redes Semánticas Naturales

De acuerdo a la definición del Modelo de Espacio Vectorial, en este modelo se realiza la búsqueda de términos en documentos; para trasladar lo anterior al campo de las Redes Semánticas Naturales, se considera que una RSN se puede interpretar como un documento, cada definidor como un término del documento y el peso del término sería el valor M del definidor, de tal manera que la ecuación 2.1 se transforma en:

$$\overrightarrow{rsn_i} = (M_{1i}, M_{2i}, M_{3i}, M_{4i}, M_{ni}) \tag{3.5}$$

En donde n es la cantidad de definidores diferentes en la red, es decir el valor J de la RSN y M representa el valor M que tiene el definidor en la RSN i .

En el MEV cuando las colecciones tienen documentos de diferente longitud, los documentos grandes tienden a tener puntuaciones más altas dado que contienen más palabras y más repeticiones de palabras, este efecto es compensado normalizando en términos del peso de las palabras en los documentos [49].

En el caso de la comparación de redes generadas con diferente cantidad de participantes es necesario normalizar el valor M de los definidores.

Para representar la información de una forma adecuada con la finalidad de realizar la comparación mediante el MEV es necesario realizar el pre-procesamiento de la información.

3.6.1. Pre-procesamiento de la Información

La normalización es una técnica de escalamiento o mapeo utilizada comúnmente en etapas de pre-procesamiento de información [79]; consiste en realizar una operación a un conjunto de valores de una magnitud dada para transformarlos en otros que estarán expresados en una escala previamente definida, comúnmente en el rango (0-1) [80].

El método de normalización que se aplica en este estudio es el denominado *min-max* que realiza una transformación lineal de los datos originales, si suponemos que \min_a y \max_a son los valores mínimos y máximos de A, se pueden obtener valores normalizados en el rango (0,1) mediante [81]:

$$v' = \frac{v - \min_a}{(\max_a - \min_a)} \quad (3.6)$$

El primer paso en el pre-procesamiento es vaciar en un vector denominado inicial[70] todos los definidores de los grupos SAM de la RSN y en otro vector se vacían los valores M correspondientes a los definidores, este vector se denomina inicialM[70], en la Figura 3.7 se presentan estos vectores para la RSN ECOLOGÍA 1, los vectores para el resto de la redes se pueden encontrar en el Apéndice A.

inicial[70]	inicialM[70]		
AGUA	49	ENFERMEDAD	76
AGUA	85	FABRICA	42
AIRE	49	FAUNA	56
AIRE	35	FLORA	75
AMBIENTE	47	GENTE	56
AMBIENTE	86	HUMO	109
AMBIENTE	66	LATA	55
AMBIENTE	59	LIMPIA	77
ANIMAL	79	MONTERREY	51
ANIMAL	96	NATURAL	92
ANIMAL	199	NATURAL	133
ARBOL	117	ORGANICO	52
ARBOL	117	OXÍGENO	86
ARBOL	136	PAPEL	100
ARBOL	270	PERSONA	137
BASURA	243	PLANTA	102
BASURA	176	PLANTA	92
BOSQUE	70	PLANTA	150
BOSQUE	126	PLANTA	171
BOTE	45	PLÁSTICO	125
BOTELLA	45	RECICLAJE	109
CALLE	49	REUSAR	202
CARRO	198	RIO	55
CARRO	66	SEMILLA	45
CASA	115	SMOG	71
CONTAMINACION	142	SUCIO	101
CONTAMINACION	57	TIENDAS	33
CUIDADO	51	TIERRA	58
DESECHO	53	TOXICO	33
ECOLOGÍA	70	TRAFICO	82
ECOLOGÍA	55	VERDE	119
ECOLOGÍA	70	VERDE	51
ECOSISTEMA	53	VERDE	115
ECOSISTEMA	96	VERDE	47
EDIFICIO	218	VIDA	63

Figura 3. 7. Formación de vectores inicial e inicialM con los definidores y los valores M de la RSN para pre-procesamiento de la información.

Para obtener el formato necesario para la comparación mediante el Modelo de Espacio Vectorial es necesario agrupar los definidores iguales y calcular una sumatoria de sus valores M para posteriormente realizar la normalización de los valores, para este proceso se desarrolló un programa en *Scilab*, el cual tiene como entrada los vectores *inicial[70]* e *inicialM[70]* y generará como salida los siguientes vectores:

- $final[j]$ Vector con los definidores sin repetición
- $finalM[j]$ Vector con la suma de los valores M de cada definidor
- $finalMNor[j]$ Vector con la suma de los valores M normalizados para cada definidor

La Figura 3.8 expone los vectores generados por el programa para la RSN ECOLOGIA 1, los vectores para el resto de los proyectos se pueden revisar en el Apéndice B.

3.7. Visualización de Redes Semánticas Naturales mediante nubes de etiquetas.

Las nubes de etiquetas (*tags clouds*) son un método conocido para la visualización de información en sitios web, representan aspectos de interés mediante la apariencia visual de las propias palabras clave, lo anterior mediante el uso de propiedades de texto tales como su tamaño, su tipo de letra y su color [82].

Wordle es una aplicación desarrollada en java que está disponible en internet, con esta aplicación se pueden generar nubes de etiquetas [83]. Esta aplicación toma como entrada un texto y genera una nube de palabras en donde el tamaño de la palabra depende de su frecuencia de aparición en el texto, también cuenta con opciones avanzadas en las que se puede proporcionar una lista de palabras y sus pesos, lo cuales representan la relevancia de la palabra en la lista, además es posible también incluir un valor hexadecimal que representa el color usado para mostrar dicha palabra.

En este trabajo se utilizó la aplicación *wordle* para realizar la representación visual de las RSN generadas; se tomó en cuenta el valor M de cada definidor en la red como el peso, de esta manera el tamaño con el que aparece el definidor en la nube está relacionado con su valor M; además se definió una escala de color para la visualización de la red, se seleccionaron cinco matices del color azul para ser utilizados en la representación, los cuales

son mostrados en la Figura 3.9, la intensidad del matiz varía dependiendo de la magnitud del valor M, a un matiz más intenso le corresponde un valor M mayor.

final[J]	finalM[J]	finalMNor[J]
AGUA	134	0.166392
AIRE	84	0.08402
AMBIENTE	258	0.370675
ANIMAL	374	0.561779
ARBOL	640	1
BASURA	419	0.635914
BOSQUE	196	0.268534
BOTE	45	0.019769
BOTELLA	45	0.019769
CALLE	49	0.026359
CARRO	264	0.38056
CASA	115	0.135091
CONTAMINACION	199	0.273476
CUIDADO	51	0.029654
DESECHO	53	0.032949
ECOLOGÍA	195	0.266886
ECOSISTEMA	149	0.191104
EDIFICIO	218	0.304778
ENFERMEDAD	76	0.07084
FABRICA	42	0.014827
FAUNA	56	0.037891
FLORA	75	0.069193
GENTE	56	0.037891
HUMO	109	0.125206
LATA	55	0.036244
LIMPIA	77	0.072488
MONTERREY	51	0.029654
NATURAL	225	0.31631
ORGANICO	52	0.031301
OXÍGENO	86	0.087315
PAPEL	100	0.110379
PERSONA	137	0.171334
PLANTA	515	0.794069
PLÁSTICO	125	0.151565
RECICLAJE	109	0.125206
REUSAR	202	0.278418
RIO	55	0.036244
SEMILLA	45	0.019769
SMOG	71	0.062603
SUCIO	101	0.112026
TIENDAS	33	0
TIERRA	58	0.041186
TOXICO	33	0
TRAFICO	82	0.080725
VERDE	332	0.492586
VIDA	63	0.049423

Figura 3. 8. Muestra la salida del programa de pre-procesamiento para obtener los vectores para la comparación.

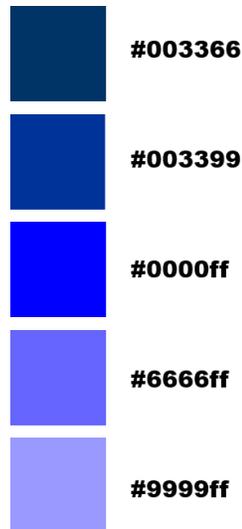


Figura 3. 9. Matices del color azul seleccionados para la representación visual.

Para asignar el matiz correspondiente a cada palabra, se determinan cinco intervalos iguales dentro del rango en el que se encuentran los Valores M de la red, a cada uno de los intervalos le corresponderá un matiz con base al valor M. El tamaño del intervalo se determina mediante la ecuación 3.7 con base al valor mínimo y máximo de M (M_{max} y M_{min}); la asignación del color se define con el algoritmo mostrado en la Figura 3.10 para la lista de definidores ordenada de manera ascendente con base al valor M.

$$inter = (M_{max} - M_{min})/5 \quad (3.7)$$

```

IF ValorM <= (Mmin + inter) THEN
  color= "9999ff"
ELSE
  IF ValorM <= (Mmin + inter*2) THEN
    color= "6666ff"
  ELSE
    IF ValorM <= (Mmin + inter*3) THEN
      color= "0000ff"
    ELSE
      IF ValorM <= (Mmin + inter*4) THEN
        color= "003399"
      ELSE
        color= "003366"
      ENDIF
    ENDIF
  ENDIF
ENDIF
ENDIF
ENDIF

```

Figura 3.10. Algoritmo para asignación de matiz de color de acuerdo al valor M para la visualización en nubes de etiquetas de RSN.

Los valores de entrada requeridos por la aplicación *wordle* para la generación de la nube de palabras para la visualización son:

- el definidor;
- el peso del definidor, que en este caso sería su Valor M;
- el código hexadecimal del matiz de color que corresponde a ese definidor de acuerdo al algoritmo en la Figura 3.10.

La Tabla 3.5 muestra la información de entrada para generar la visualización de la RSN ECOLOGÍA 1.

La Figura 3.11 muestra un ejemplo de la visualización de nube de etiquetas para ECOLOGÍA 1.

Tabla 3. 5. Información requerida para generar la visualización de RSN ECOLOGÍA 1 mediante *wordle*.

Definidor	Valor M	Color	Definidor	Valor M	Color
Arbol	640	003366	Aire	84	9999ff
Planta	515	003399	Tráfico	82	9999ff
Basura	419	003399	Limpia	77	9999ff
Animal	374	0000ff	Enfermedad	76	9999ff
Verde	332	0000ff	Flora	75	9999ff
Carro	264	6666ff	Smog	71	9999ff
Ambiente	258	6666ff	Metrópolis	66	9999ff
Natural	225	6666ff	Vida	63	9999ff
Edificio	218	6666ff	Tierra	58	9999ff
Reusar	202	6666ff	Fauna	56	9999ff
Contaminación	199	6666ff	Gente	56	9999ff
Bosque	196	6666ff	Lata	55	9999ff
Ecología	195	6666ff	Río	55	9999ff
Ecosistema	149	9999ff	Desecho	53	9999ff
Persona	137	9999ff	Orgánico	52	9999ff
Agua	134	9999ff	Cuidado	51	9999ff
Plástico	125	9999ff	Calle	49	9999ff
Casa	115	9999ff	Bote	45	9999ff
Humo	109	9999ff	Botella	45	9999ff
Reciclaje	109	9999ff	Semilla	45	9999ff
Sucio	101	9999ff	Fábrica	42	9999ff
Papel	100	9999ff	Tienda	33	9999ff
Oxígeno	86	9999ff	Tóxico	33	9999ff

Al trasladar el proceso que sigue el Modelo de Espacio Vectorial hacia la comparación entre Redes Semánticas Naturales, se considera que la comparación se realizaría entre dos redes semánticas en donde los definidores de los grupos SAM se compararán y el peso de los términos estaría dado por los valores M normalizados de los definidores, las técnicas de comparación son las indicadas para MEV y el resultado de la similitud sería interpretado como el parecido que existe entre las redes comparadas; la analogía anterior es representada en la Figura 3.12.

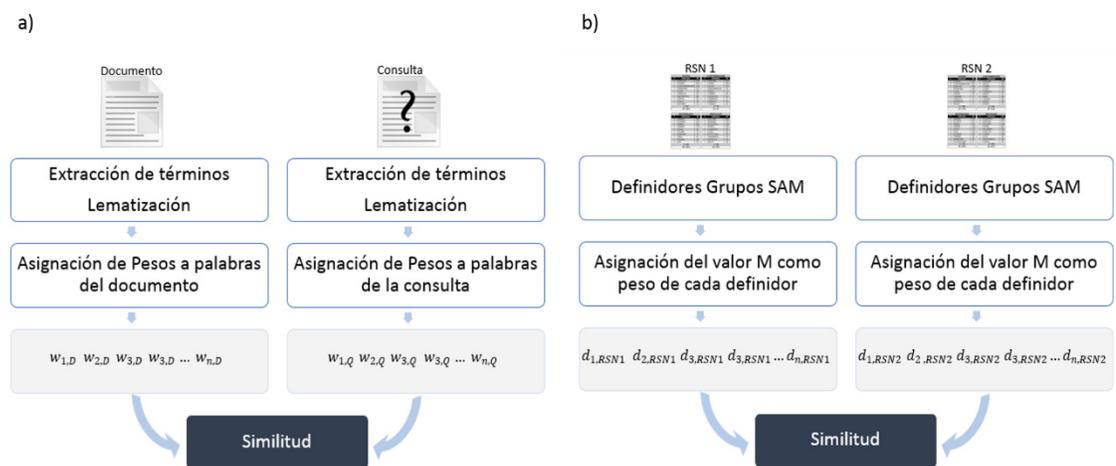


Figura 3.12. Representa la analogía entre el proceso del Modelo de Espacio Vectorial clásico de Recuperación de Información y el proceso del Modelo de Espacio Vectorial aplicado a la comparación de RSN. En a) se muestra el proceso de comparación entre un documento y una consulta y en b) se muestra el proceso de comparación entre dos RSN.

3.8.1. Proceso de comparación de RSN mediante MEV

Para realizar la comparación entre dos RSN se estructura una matriz en donde las columnas representan los definidores y las filas las RSN a comparar; cada elemento representa el Valor M normalizado del definidor j en la RSN i , como se muestra en la Figura 3.13.

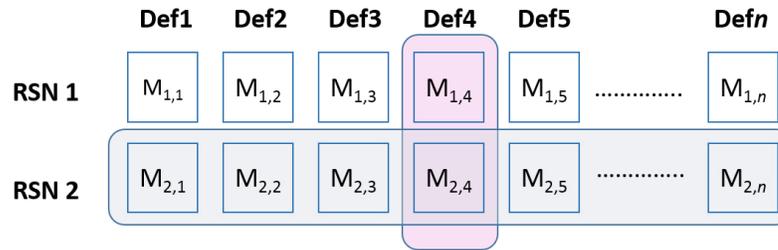


Figura 3. 13. En esta matriz cada fila representa una RSN, cada columna representa un definidor y cada elemento representa el Valor M de cada definidor en la RSN correspondiente.

Se realiza la comparación entre las RSN mediante las medidas de similitud definidas para el Modelo de Espacio Vectorial que son la similitud cosenoidal, el coeficiente de Jaccard y el coeficiente de Dice calculados aplicando las ecuaciones 2.3, 2.4 y 2.5 respectivamente.

Se desarrolló un programa en *Scilab* para realizar el proceso de comparación, este programa tiene como entrada los vectores obtenidos en el pre-procesamiento de la información para cada RSN a comparar, los cuales son el vector de definidores de la RSN $final[j]$ y el vector de Valor M normalizado para cada definidor $finalMNor[j]$, estos son mostrados en la Figura 3.8. El programa tiene como salida una matriz denominada $matrizcomp[2,j]$, la cual contiene los definidores y los valores M normalizados correspondientes por definidor, a aquellos definidores que sólo existen en una de las redes se les asigna un valor M de cero en el espacio del definidor correspondiente y son mostrados con espacios vacíos. La Tabla 3.6 muestra el ejemplo de la matriz de salida para los primeros diez definidores en la comparación entre dos RSN.

Tabla 3.6. Ejemplo de matriz de salida para la comparación de RSN mediante MEV, los espacios vacíos indican que en esa red no existió ese definidor.

	AGUA	AIRE	ALUMINIO	AMBIENTE	ANIMAL	ARBOL	BASURA	BOSQUE	BOTE	BOTELLA
PROYECTO 1	0.166	0.084		0.371	0.562	1.000	0.636	0.269	0.020	0.020
PROYECTO 2	0.359	0.181	0.139	0.213	0.840	0.965	0.518	0.174		

Los Valores M normalizados con un valor de 1 representan que el valor máximo de M ocurrió en tal definidor, asimismo, los valores M normalizados con valor 0 representan que el valor mínimo de M ocurrió en ese definidor, dado que la representación del 0 normalizado indica un valor mínimo de M, en los definidores con valor 0 se representó el valor 0.001.

Una vez que se tiene definida la matriz de salida *matrizcomp*[2,j], se realizan las operaciones de cálculo de similitud para obtener el parecido entre estas dos redes. La Figura 3.14 expone la sección del código del programa en la que se calculan los coeficientes de similitud.

```
//Cosenoidal
sumsup = 0
sum1inf = 0
sum2inf = 0
for j=1: tam
    sumsup = sumsup + (matrizcomp (1 , j) * matrizcomp (2 , j))
    sum1inf = sum1inf + matrizcomp (1 , j) ** 2
    sum2inf=sum2inf+matrizcomp (2 , j) ** 2
end
raiz1 = sqrt (sum1inf)
raiz2 = sqrt (sum2inf)
cosenoidal=sumsup / (raiz1*raiz2)
printf("cosenoidal %1.3f\n", cosenoidal)

// Coeficiente de Dice
dice=2 * sumsup / (sum1inf + sum2inf)
printf ("dice %1.3f\n", dice)

//Coeficiente de Jaccard
jaccard = sumsup / (sum1inf + sum2inf - sumsup)
printf ("jaccard %1.3f\n", jaccard)
```

Figura 3.14. Representa el cálculo de la similitud cosenoidal, y los coeficientes de Dice y Jaccard para la comparación de las RSN.

En este trabajo se realizaron las siguientes comparaciones:

- Entre las RSN de principiantes;
- Entre las RSN de expertos;
- Entre la Red Total de Principiantes con la red de cada experto;
- Entre la Red Total de Expertos con cada RSN de principiantes.
- Entre la Red Total de Principiantes con la Red Total de Expertos.

Los resultados obtenidos de estas comparaciones fueron estructurados en tablas y además fueron analizados, en la siguiente sección se exponen los resultados obtenidos.

CAPÍTULO 4

RESULTADOS Y DISCUSIÓN

4.1. Resultados de la generación de RSN para principiantes

En esta sección se exponen las RSN de la temática de ecología que fueron generadas mediante la plataforma *RSNweb* para los cuatro grupos de principiantes, las cuales son mostradas de la Figura 4.1 a la Figura 4.4.

PROYECTO - ECOLOGIA 1

CIUDAD			CONTAMINACION			ECOLOGIA			MEDIO AMBIENTE		
F	Definidor	M	F	Definidor	M	F	Definidor	M	F	Definidor	M
1	EDIFICIO	218	2	BASURA	243	4	VERDE	119	2	NATURAL	133
2	CARRO	198	1	HUMO	109	4	ARBOL	117	4	ARBOL	117
2	CONTAMINACIÓN	142	1	SUCIO	101	1	RECICLAJE	109	3	ANIMAL	96
1	PERSONA	137	1	ENFERMEDAD	76	4	PLANTA	102	2	ECOSISTEMA	96
1	CASA	115	1	SMOG	71	2	NATURAL	92	4	PLANTA	92
1	TRAFICO	82	2	CARRO	66	4	AMBIENTE	86	3	ECOLOGIA	70
1	METROPOLI	66	1	DESECHO	53	3	ANIMAL	79	2	CONTAMINACIÓN	57
1	GENTE	56	4	AMBIENTE	47	1	LIMPIA	77	4	VERDE	51
1	CALLE	49	1	FABRICA	42	2	ECOSISTEMA	53	2	AGUA	49
1	TIENDA	33	1	TOXICO	33	1	CUIDADO	51	2	AIRE	49
J = 97			J = 79			J = 66			J = 85		
G = 18.5			G = 21			G = 6.8			G = 8.4		

NATURALEZA			RECICLAJE			REFORESTACION		
F	Definidor	M	F	Definidor	M	F	Definidor	M
3	ANIMAL	199	1	REUSAR	202	4	ARBOL	270
4	PLANTA	150	2	BASURA	176	4	PLANTA	171
4	ARBOL	136	1	PLÁSTICO	125	2	BOSQUE	126
4	VERDE	115	1	PAPEL	100	1	OXÍGENO	86
2	AGUA	85	4	AMBIENTE	66	3	ECOLOGIA	70
1	FLORA	75	1	LATA	55	4	AMBIENTE	59
2	BOSQUE	70	3	ECOLOGIA	55	1	TIERRA	58
1	VIDA	63	1	ORGANICO	52	4	VERDE	47
1	FAUNA	56	1	BOTELLA	45	1	SEMILLA	45
1	RIO	55	1	BOTE	45	2	AIRE	35
J = 85			J = 59			J = 73		
G = 14.4			G = 15.7			G = 23.5		

Figura 4. 1. Representación de grupos SAM de RSN del proyecto ECOLOGÍA 1.

PROYECTO – ECOLOGIA 2

CIUDAD			CONTAMINACION			ECOLOGIA			MEDIO AMBIENTE		
F	Definidor	M	F	Definidor	M	F	Definidor	M	F	Definidor	M
2	CARRO	247	2	BASURA	246	4	PLANTA	214	4	ANIMAL	210
2	CONTAMINACIÓN	230	1	SUCIO	162	3	NATURAL	201	4	ARBOL	176
1	EDIFICIO	204	2	CARRO	143	1	RECICLAJE	181	4	PLANTA	155
1	PERSONA	195	1	HUMO	126	2	AMBIENTE	166	3	NATURAL	134
1	CASA	133	1	FABRICA	93	4	VERDE	139	2	AIRE	129
1	GENTE	93	1	SMOG	92	4	ANIMAL	135	2	ECOSISTEMA	114
1	GRANDE	77	2	AIRE	90	1	CUIDADO	108	4	AGUA	105
1	TRAFICO	72	4	TIERRA	66	4	ARBOL	102	2	CONTAMINACIÓN	99
1	EMPRESA	72	4	AGUA	59	2	LIMPIA	75	4	VERDE	87
2	RUIDO	71	2	RUIDO	53	4	AGUA	75	4	TIERRA	82
J = 158			J = 119			J = 96			J = 91		
G = 17.6			G = 19.3			G = 13.9			G = 12.8		

NATURALEZA			RECICLAJE			REFORESTACION		
F	Definidor	M	F	Definidor	M	F	Definidor	M
4	ANIMAL	351	1	PLASTICO	296	4	ARBOL	396
4	PLANTA	301	2	BASURA	270	4	PLANTA	272
4	ARBOL	237	1	REUSAR	228	2	BOSQUE	114
4	VERDE	156	1	PAPEL	207	4	ANIMAL	105
4	AGUA	137	1	ALUMINIO	182	4	VERDE	86
2	VIDA	102	1	CARTON	140	1	OXIGENO	83
2	BOSQUE	99	2	LIMPIA	108	2	VIDA	72
4	TIERRA	83	1	LATA	95	3	NATURAL	62
2	AMBIENTE	81	1	INORGANICO	86	4	TIERRA	61
2	ECOSISTEMA	77	1	VIDRIO	85	1	TALA	59
J = 122			J = 102			J = 110		
G = 27.4			G = 21.1			G = 33.7		

Figura 4. 2. Representación de grupos SAM de RSN del proyecto ECOLOGÍA 2.

PROYECTO - ECOLOGIA 3

CIUDAD			CONTAMINACION			ECOLOGIA			MEDIO AMBIENTE		
F	Definidor	M	F	Definidor	M	F	Definidor	M	F	Definidor	M
4	CONTAMINACIÓN	172	2	BASURA	255	3	NATURAL	151	3	ANIMAL	153
2	CARRO	161	1	HUMO	180	3	AMBIENTE	124	4	CONTAMINACION	99
1	EDIFICIO	155	2	CARRO	108	3	VERDE	116	3	NATURAL	93
1	GENTE	99	1	FABRICA	105	1	RECICLAJE	114	4	ARBOL	73
1	PERSONA	94	1	SMOG	90	4	ARBOL	106	2	ECOSISTEMA	73
1	RUIDO	66	1	SUCIO	73	3	ANIMAL	80	3	AIRE	58
1	CASA	62	1	ENFERMEDAD	72	3	CUIDADO	55	4	PLANTA	57
1	EMPRESA	52	3	AIRE	67	4	PLANTA	54	1	CLIMA	52
1	GRANDE	49	1	AGUA	53	1	MEDIO	52	3	CUIDADO	50
1	TRAFICO	47	1	CIUDAD	51	4	CONTAMINAC...	48	2	OXIGENO	47
J = 116			J = 87			J = 72			J = 87		
G = 12.5			G = 20.4			G = 10.3			G = 10.6		

NATURALEZA			RECICLAJE			REFORESTACION		
F	Definidor	M	F	Definidor	M	F	Definidor	M
3	ANIMAL	226	2	BASURA	172	4	ARBOL	273
4	PLANTA	165	1	REUSAR	111	4	PLANTA	113
4	ARBOL	140	1	PLÁSTICO	108	2	OXIGENO	97
2	VIDA	86	1	PAPEL	83	1	BOSQUE	80
3	AMBIENTE	84	1	AYUDA	72	2	VIDA	62
3	VERDE	78	1	LATA	51	3	AMBIENTE	48
1	MONTAÑA	69	3	NATURAL	50	3	AIRE	40
2	ECOSISTEMA	65	3	CUIDADO	44	1	TALA	39
1	FAUNA	55	4	CONTAMINACIÓN	42	1	NECESIDAD	35
1	RIO	50	1	ECOLOGIA	37	3	VERDE	34
J = 90			J = 73			J = 91		
G = 17.6			G = 13.5			G = 23.9		

Figura 4. 3. Representación de grupos SAM de RSN del proyecto ECOLOGÍA 3.

PROYECTO - ECOLOGIA 4

CIUDAD			CONTAMINACION			ECOLOGIA			MEDIO AMBIENTE		
F	Definidor	M	F	Definidor	M	F	Definidor	M	F	Definidor	M
1	EDIFICIO	288	2	BASURA	268	3	NATURAL	224	4	ANIMAL	191
2	CARRO	249	1	SUCIO	115	4	ARBOL	162	4	ARBOL	183
3	CONTAMINACION	233	1	CIUDAD	104	1	RECICLAJE	140	3	NATURAL	160
1	PERSONA	148	1	SMOG	92	1	CUIDADO	109	4	AGUA	121
1	POBLACION	107	1	HUMO	91	4	PLANTA	107	3	CONTAMINACION	119
1	CASA	93	3	AIRE	91	3	AMBIENTE	104	4	PLANTA	106
1	GRANDE	73	2	CARRO	90	3	ECOSISTEMA	92	3	AIRE	103
1	RUIDO	68	4	AGUA	81	3	VERDE	90	2	TIERRA	90
1	CALLE	65	1	FABRICA	78	4	ANIMAL	84	3	ECOSISTEMA	80
1	GENTE	55	1	ENFERMEDAD	73	1	LIMPIA	60	3	VIDA	72
J = 109			J = 100			J = 78			J = 90		
G = 23.3			G = 19.5			G = 16.4			G = 11.9		

NATURALEZA			RECICLAJE			REFORESTACION		
F	Definidor	M	F	Definidor	M	F	Definidor	M
4	ANIMAL	301	1	REUSAR	179	4	ARBOL	411
4	ARBOL	280	1	PLASTICO	170	4	PLANTA	258
4	PLANTA	192	1	PAPEL	163	1	OXIGENO	117
4	AGUA	171	2	BASURA	156	1	BOSQUE	102
3	VIDA	111	1	ALUMINIO	107	3	NATURAL	93
3	VERDE	104	3	AMBIENTE	85	3	VIDA	73
3	ECOSISTEMA	100	1	ECOLOGIA	81	4	ANIMAL	64
2	TIERRA	97	1	BOTELLA	75	3	VERDE	59
3	AIRE	88	3	CONTAMINAC...	66	3	AMBIENTE	54
1	FLORA	65	1	MEDIO	65	1	TALA	51
J = 98			J = 89			J = 96		
G = 23.6			G = 11.4			G = 36		

Figura 4. 4. Representación de grupos SAM de RSN del proyecto ECOLOGÍA 4.

4.1.1. Generación de Red Total de Principiantes

Se generó una RSN tomando en cuenta los definidores de las cuatro pruebas (ECOLOGÍA 1, ECOLOGÍA 2, ECOLOGÍA 3, ECOLOGÍA 4), es decir como si fuera un sólo grupo de 159 participantes, a esta red se le denominó Red Total de Principiantes, se formaron los grupos SAM y se realizó el cálculo del valor J, valor M y valor G. La Figura 4.5 presenta la información generada de la Red Total de Principiantes, esta red se creó con el objetivo de compararla posteriormente con la red de cada uno de los expertos así como con la Red Total de Expertos.

PROYECTO RED TOTAL

CIUDAD			CONTAMINACION			ECOLOGIA			MEDIO AMBIENTE		
F	Definidor	M	F	Definidor	M	F	Definidor	M	F	Definidor	M
1	EDIFICIO	865	2	BASURA	1012	3	NATURAL	668	4	ANIMAL	650
2	CARRO	855	1	HUMO	506	1	RECICLAJE	544	4	ARBOL	549
2	CONTAMINACIÓN	777	1	SUCIO	451	4	ARBOL	487	3	NATURAL	520
1	PERSONA	574	2	CARRO	407	4	AMBIENTE	480	4	PLANTA	410
1	CASA	403	1	SMOG	345	4	PLANTA	477	2	CONTAMINACION	374
1	GENTE	303	1	FABRICA	318	3	VERDE	464	3	ECOSISTEMA	363
1	TRAFICO	253	2	AIRE	279	4	ANIMAL	378	2	AIRE	339
1	GRANDE	229	1	ENFERMEDAD	254	2	CUIDADO	323	3	AGUA	300
1	RUIDO	229	1	CIUDAD	225	2	LIMPIA	241	2	ECOLOGIA	246
1	CALLE	188	3	AGUA	198	3	ECOSISTEMA	187	2	CUIDADO	219
J = 270			J = 214			J = 168			J = 180		
G = 67.7			G = 81.4			G = 48.1			G = 43.1		

NATURALEZA			RECICLAJE			REFORESTACION		
F	Definidor	M	F	Definidor	M	F	Definidor	M
4	ANIMAL	1077	2	BASURA	774	4	ARBOL	1350
4	PLANTA	808	1	REUSAR	720	4	PLANTA	814
4	ARBOL	793	1	PLASTICO	699	2	BOSQUE	422
3	VERDE	453	1	PAPEL	553	1	OXIGENO	383
3	AGUA	429	1	ALUMINIO	339	2	VIDA	240
2	VIDA	362	4	AMBIENTE	266	3	VERDE	226
2	BOSQUE	271	2	ECOLOGIA	251	4	AMBIENTE	203
3	ECOSISTEMA	261	1	LATA	241	4	ANIMAL	201
4	AMBIENTE	249	1	CARTON	238	3	NATURAL	198
1	FLORA	246	2	LIMPIA	212	1	TALA	158
J = 221			J = 182			J = 224		
G = 83.1			G = 56.2			G = 119.2		

Figura 4. 5. Representación de la RED TOTAL formada con la información obtenida de todos los principiantes.

4.2. Resultados de la generación de Red de Expertos

En esta sección se muestran las plantillas realizadas por los cuatro expertos en ecología, dichas representaciones se pueden observar de la Figura 4.6 a la Figura 4.9

PLANTILLA EXPERTO 1

CIUDAD			CONTAMINACION			ECOLOGIA			MEDIO AMBIENTE		
F	Definidor	M	F	Definidor	M	F	Definidor	M	F	Definidor	M
1	ESTRES	5	1	HUMO	7	1	ORGANISMO	9	2	INTERACCIÓN	10
2	GENTE	7	1	BASURA	5	2	INTERACCIÓN	7	3	VERDE	2
2	CASA	8	1	DESCOMPOSICIÓN	9	1	NATURALEZA	8	1	FLUJO	8
1	CEMENTO	6	1	DETERIORO	10	1	MAR	3	2	CICLO	7
1	METRÓPOLI	9	1	QUÍMICOS	8	2	TIERRA	4	2	BIODIVERSIDAD	6
1	POBLACIÓN	10	1	SMOG	6	2	BIODIVERSIDAD	10	1	RELACIONES	4
1	CARRO	4	1	CONTAGIO	2	1	VIENTO	2	2	GENTE	1
2	FORTALEZA	2	1	MEZCLA	3	2	CASA	1	2	RECURSOS	5
1	PANORÁMICO	3	1	INFECCIÓN	1	1	SISTEMA	5	1	HÁBITAT	3
1	SEÑALIZACIÓN	1	1	PUDRICIÓN	4	2	RECURSOS	6	2	UNIVERSO	9

NATURALEZA			RECICLAJE			REFORESTACION			ENERGIA		
F	Definidor	M	F	Definidor	M	F	Definidor	M	F	Definidor	M
1	VIDA	10	1	TRANSFORMAR	9	1	PLANTA	9	1	SOL	8
3	VERDE	7	1	PROCESAR	10	3	VERDE	8	1	POTENCIA	5
2	UNIVERSO	9	1	APROVECHAR	3	1	SIERRA	5	1	INTENSIDAD	4
2	TIERRA	8	1	RETOMAR	4	1	PODAS	3	1	IMPULSO	3
1	ESTRUCTURA	6	1	REHABILITAR	7	1	INJERTO	2	2	FORTALEZA	9
1	ESENCIA	2	1	ENCAUSAR	2	1	TRANQUILIDAD	1	1	FUERTE	1
1	IDENTIDAD	1	1	RECUPERAR	8	1	ARBOLAR	6	1	PODER	7
1	VARIEDAD	5	1	REUSAR	6	1	BOSQUE	7	1	LUZ	6
1	ARMONÍA	3	1	REDUCIR	5	1	HOJA	4	1	NUCLEO	2
2	CICLO	4	1	COMPOSTA	1	2	SUSTENTABILIDAD	10	2	SUSTENTABILIDAD	10

Figura 4. 6. Muestra la Plantilla que representa la RSN del Experto 1.

PLANTILLA EXPERTO 2

CIUDAD			CONTAMINACION			ECOLOGIA			MEDIO AMBIENTE		
F	Definidor	M	F	Definidor	M	F	Definidor	M	F	Definidor	M
1	CASA	6	1	INDICES	1	1	CONSERVACIÓN	7	1	CONTAMINANTE	10
1	CARRO	7	1	GAS	10	2	BOSQUE	6	1	CUIDADO	9
2	POBLACIÓN	9	2	ÁRBOL	2	1	NICHO	8	1	AIRE	6
2	PARQUE	3	1	NINO	3	1	PLANETA	9	1	PUREZA	5
1	CALLE	1	1	OXÍGENO	9	1	ESTUDIO	5	1	OCÉANO	4
2	FABRICA	8	1	RIESGO	8	1	DESIERTO	4	1	BIÓSFERA	7
1	RESTAURANTE	2	1	ENFERMEDAD	7	1	RELACIÓN	1	2	DESTRUCCIÓN	8
1	EMPRESA	10	2	FÁBRICA	6	5	HUMANO	2	2	POBLACIÓN	3
1	CONTAMINACIÓN	4	1	TRANSPORTE	5	2	EQUILIBRIO	10	4	CONCIENCIA	1
1	EDIFICIO	5	2	AGUA	4	1	INTERACCIÓN	3	5	HUMANO	2

NATURALEZA			RECICLAJE			REFORESTACION			ENERGIA		
F	Definidor	M	F	Definidor	M	F	Definidor	M	F	Definidor	M
1	VERDE	2	1	BASURA	6	2	ÁRBOL	10	1	SOLAR	8
1	FAUNA	3	1	CREATIVIDAD	5	1	ESPACIO	8	1	UTILIZACIÓN	10
1	FLORA	4	4	CONCIENCIA	10	1	SUELO	9	1	EÓLICA	6
1	PAISAJE	5	1	DESECHO	9	1	VEGETACIÓN	7	1	NUCLEAR	7
1	BELLEZA	6	1	TÓXICO	8	1	NUTRIENTES	1	1	HIDROELÉCTRICA	5
1	ARMONÍA	7	1	APROVECHAMIENTO	7	2	PARQUE	2	1	HIDROCARBUROS	4
2	EQUILIBRIO	10	1	PLÁSTICO	2	1	RURAL	3	4	CONCIENCIA	9
2	DESTRUCCIÓN	9	2	AGUA	4	2	BOSQUE	6	5	HUMANO	3
4	CONCIENCIA	8	1	CONTENEDOR	3	1	REGENERACIÓN	5	1	POLÍTICA	1
5	HUMANO	1	1	ALUMINIO	1	5	HUMANO	4	1	COMBUSTIBLE	2

Figura 4. 7. Muestra la Plantilla que representa la RSN del Experto 2.

PLANTILLA EXPERTO 3

CIUDAD			CONTAMINACION			ECOLOGIA			MEDIO AMBIENTE		
F	Definidor	M	F	Definidor	M	F	Definidor	M	F	Definidor	M
1	URBANO	10	1	DANO	1	1	BALANCE	6	1	HABITAT	10
1	METRÓPOLI	9	1	AIRE	10	1	CASA	9	2	RELACIÓN	9
1	TRÁFICO	8	1	PROBLEMÁTICA	3	1	CIENCIA	10	1	BIÓSFERA	8
1	ARTIFICIAL	5	1	SMOG	9	1	INVESTIGACIÓN	5	1	SISTEMA	3
1	URBANO	1	1	PARTICULAR	7	1	HOMEÓSTASIS	2	1	SOCIAL	7
1	EDIFICIO	7	1	RUIDO	6	2	RELACIÓN	1	1	BIÓTICO	5
1	GENTE	3	1	LETRERO	4	1	PLANETA	8	1	ABIÓTICO	4
1	LUZ	4	2	BASURA	8	1	ESTUDIO	7	1	CLIMA	6
1	ENERGÍA	2	1	TÓXICO	2	1	VIDA	4	1	OCÉANO	1
1	POBLACIÓN	6	1	GRIS	5	1	ARMONÍA	3	1	MONTANAS	2

NATURALEZA			RECICLAJE			REFORESTACION			ENERGÍA		
F	Definidor	M	F	Definidor	M	F	Definidor	M	F	Definidor	M
1	FLORA	10	1	UTILIDAD	8	1	VERDE	6	1	EÓLICA	6
1	FAUNA	9	1	REUTILIZAR	10	1	ÁRBOL	10	1	SOLAR	10
1	VIDA	6	1	RESIDUO	9	2	VEGETACIÓN	9	1	ELECTRICIDAD	9
1	ORGANISMO	3	2	BASURA	7	2	RECUPERAR	5	1	COMBUSTIBLE	4
1	COLOR	8	1	SUBPRODUCTO	6	1	PLANTA	8	1	NUCLEAR	8
1	PAISAJE	7	2	RECUPERAR	5	1	ÁREA VERDE	3	1	LUZ	7
1	HOMBRE	5	1	REEVALUAR	4	1	PARQUE	4	1	GAS	1
1	ECOSISTEMA	4	1	PLÁSTICO	3	1	SUELO	2	1	HIDROCARBURO	3
1	ANIMAL	2	1	AGUA	1	1	BOSQUE	7	1	TERMOELÉCTRICA	5
2	VEGETACIÓN	1	1	CONSUMISMO	2	1	REGENERACIÓN	1	1	BATERÍA	2

Figura 4. 8. Muestra la Plantilla que representa la RSN del Experto 3.

PLANTILLA EXPERTO 4

CIUDAD			CONTAMINACION			ECOLOGIA			MEDIO AMBIENTE		
F	Definidor	M	F	Definidor	M	F	Definidor	M	F	Definidor	M
2	CARRO	8	2	CARRO	10	1	AREA VERDE	10	2	CONTAMINACION	10
1	AVENIDA	5	1	GASOLINERÍA	2	2	CONTAMINACIÓN	8	1	OZONO	9
1	EDIFICIO	10	1	INDUSTRIA	9	2	HOMBRE	9	1	INVERNADERO	8
1	NEGOCIO	4	1	SMOG	8	2	BASURA	3	3	SUELO	3
2	POBLACIÓN	9	1	RESTAURANTE	4	1	DESECHO	4	4	AGUA	4
1	TRANSPORTE	6	1	CONSTRUCCIÓN	5	2	GAS	7	1	PLANTA	5
1	AMBULANTE	3	1	ALCANTARILLA	3	1	AZUFRE	1	2	DESECHO	2
1	SEMÁFORO	1	2	BASURA	7	1	COMUNIDAD	2	2	POBLACION	7
1	VIALIDAD	7	1	AGUAS RESIDUALES	6	1	SUELO	6	1	URBANIZACION	6
1	SENALAMIENTO	2	1	AMBULANTE	1	4	AGUA	5	1	NEGOCIO	1

NATURALEZA			RECICLAJE			REFORESTACION			ENERGÍA		
F	Definidor	M	F	Definidor	M	F	Definidor	M	F	Definidor	M
1	BIODIVERSIDAD	10	1	VIDRIO	2	1	VEGETACION	10	1	SOLAR	10
2	HOMBRE	9	1	CARTON	5	3	SUELO	8	1	NUCLEAR	1
2	GAS	3	1	PAPEL	4	4	AGUA	9	1	HIDROELECTRICA	9
1	PRODUCTIVIDAD	4	1	ALUMINIO	3	1	NUTRIENTE	6	1	GEOTERMICA	2
1	SOL	7	1	FIERRO	1	1	ABONO	5	4	AGUA	8
1	EROSION	1	1	MEDIO AMBIENTE	6	1	ESPECIES	1	3	SUELO	7
1	ENFERMEDAD	2	1	DESARROLLO	7	1	RURAL	7	1	LUZ	6
2	BASURA	5	1	SUSTENTABILIDAD	10	1	COMPOSTA	4	1	ELECTRODOMESTICO	3
1	CLIMA	8	1	APROVECHAMIENTO	9	1	HUERTO	3	1	GAS	4
1	LLUVIA	6	1	RECURSOS	8	1	FORRAJE	2	1	ALIMENTO	5

Figura 4. 9. Muestra la Plantilla que representa la RSN del Experto 4.

Se generó una RSN con la información obtenida de todos los expertos, esta RSN se denominó Red Total de Expertos, se formaron los grupos SAM con los mayores valores de M y se calcularon los indicadores correspondientes, la Figura 4.10 muestra la representación de la Red Total de Expertos, la cual se comparará con la Red Total de principiantes.

RED EXPERTOS

CIUDAD			CONTAMINACION			ECOLOGIA			MEDIO AMBIENTE		
F	Definidor	M	F	Definidor	M	F	Definidor	M	F	Definidor	M
2	POBLACION	34	1	SMOG	23	1	PLANETA	17	1	BIOSFERA	15
1	EDIFICIO	22	2	BASURA	20	1	ESTUDIO	12	1	RELACION	13
2	CARRO	19	1	AIRE	10	1	AREA VERDE	10	1	HABITAT	13
1	METROPOLI	18	2	CARRO	10	2	BIODIVERSIDAD	10	1	CONTAMINACION	10
2	CASA	14	1	DETERIORO	10	2	CASA	10	1	CONTAMINANTES	10
1	URBANO	11	1	GAS	10	1	CIENCIA	10	2	INTERACCION	10
1	EMPRESA	10	1	DESCOMPOSICION	9	2	EQUILIBRIO	10	2	POBLACION	10
1	GENTE	10	1	INDUSTRIA	9	2	INTERACCION	10	1	CUIDADO	9
1	FABRICA	8	1	OXIGENO	9	2	HOMBRE	9	1	OZONO	9
1	TRAFICO	8	1	QUIMICOS	8	1	ORGANISMO	9	1	UNIVERSO	9
J = 28 G = 2.6			J = 36 G = 1.5			J = 34 G = 0.8			J = 35 G = 0.6		
NATURALEZA			RECICLAJE			REFORESTACION			ENERGIA		
F	Definidor	M	F	Definidor	M	F	Definidor	M	F	Definidor	M
1	VIDA	16	1	APROVECHAR	19	1	VEGETACION	26	1	SOLAR	28
1	FLORA	14	2	BASURA	13	1	ARBOL	20	1	LUZ	19
2	HOMBRE	14	1	RECUPERAR	13	1	BOSQUE	20	1	NUCLEAR	16
1	FAUNA	12	2	CONCIENCIA	10	1	SUELO	19	1	HIDROELECTRICA	14
1	PAISAJE	12	1	PROCESAR	10	1	PLANTA	17	1	EOLICA	12
1	ARMONIA	10	1	REUTILIZAR	10	2	VERDE	14	3	SUSTENTABILIDAD	10
2	BIODIVERSIDAD	10	3	SUSTENTABILIDAD	10	1	RURAL	10	1	UTILIZACION	10
2	EQUILIBRIO	10	1	DESECHO	9	3	SUSTENTABILIDAD	10	2	CONCIENCIA	9
1	DESTRUCCION	9	1	RESIDUO	9	1	AGUA	9	1	ELECTRICIDAD	9
2	VERDE	9	1	TRANSFORMACION	9	1	GENERADOR	8	1	FORTALEZA	9
J = 33 G = 0.7			J = 33 G = 1			J = 28 G = 1.9			J = 29 G = 1.9		

Figura 4. 10. Representación de Red Total de Expertos, generada con la información de los cuatro expertos participantes.

4.3. Representación visual de la RSN generadas

A continuación se muestran las visualizaciones realizadas para las diferentes redes generadas mediante la aplicación *wordle* de acuerdo a las especificaciones definidas en el apartado 3.7.



Figura 4. 18. Representación visual mediante nubes de etiquetas de la plantilla de experto 4.

4.3.3. Representación visual de Redes Totales

Las Red Total formada con los definidores de todos los principiantes se puede observar en la representación visual de la Figura 4.19, así como la Red Total de Expertos compuesta de los definidores de todos los expertos.



Figura 4. 19. Visualización de Red Total de Principiantes y de Red Total de Expertos mediante nube de etiquetas.

4.4. Resultado de la comparación de Redes Semánticas Naturales mediante Modelo de Espacio Vectorial

Se realizaron diferentes comparaciones entre las RSN generadas aplicando la adaptación del Modelo de Espacio Vectorial definida en la sección de Metodología, en las siguientes subsecciones se muestran los resultados obtenidos.

4.4.1. Comparación entre RSN de principiantes

Aplicando la metodología definida en el apartado 3.8, se ejecutó el programa para la generación de la matriz de comparación entre las cuatro RSN de principiantes generadas; en la Tabla 4.1 se muestran las matrices de comparación obtenidas entre la RSN de ECOLOGÍA 1 con la RSN de ECOLOGÍA 2, así como la comparación de la RSN de ECOLOGÍA 1 con la RSN de ECOLOGÍA 4. Como se describió antes, el espacio vacío significa que el definidor no aparece en esa red pero sí existe en la red con la que se compara.

En el Apéndice C se pueden consultar las matrices de comparación generadas para el resto de las comparaciones entre principiantes.

Tabla 4. 1. Muestra la matriz de comparación entre la RSN ECOLOGÍA 1 con la RSN ECOLOGÍA 2 y de la RSN ECOLOGÍA1 con la RSN ECOLOGÍA 4.

	ECO 1	ECO 2		ECO 1	ECO 4
AGUA	0.166	0.359	AGUA	0.166	0.327
AIRE	0.084	0.181	AIRE	0.084	0.235
ALUMINIO		0.139	ALUMINIO		0.057
AMBIENTE	0.371	0.213	AMBIENTE	0.371	0.195
ANIMAL	0.562	0.840	ANIMAL	0.562	0.598
ARBOL	1.000	0.965	ARBOL	1.000	1.000
BASURA	0.636	0.518	BASURA	0.636	0.379
BOSQUE	0.269	0.174	BOSQUE	0.269	0.052
BOTE	0.020		BOTE	0.020	
BOTELLA	0.020		BOTELLA	0.020	0.024
CALLE	0.026		CALLE	0.026	0.014
CARRO	0.381	0.375	CARRO	0.381	0.292
CARTÓN		0.092	CASA	0.135	0.043
CASA	0.135	0.084	CIUDAD		0.054
CONTAMINAC...	0.273	0.306	CONTAMINAC...	0.273	0.373
CUIDADO	0.030	0.055	CUIDADO	0.030	0.059
DESECHO	0.033		DESECHO	0.033	
ECOLOGÍA	0.267		ECOLOGÍA	0.267	0.030
ECOSISTEMA	0.191	0.149	ECOSISTEMA	0.191	0.224
EDIFICIO	0.305	0.164	EDIFICIO	0.305	0.241
EMPRESA		0.015	ENFERMEDAD	0.071	0.022
ENFERMEDAD	0.071		FABRICA	0.015	0.027
FABRICA	0.015	0.039	FAUNA	0.038	
FAUNA	0.038		FLORA	0.069	0.014
FLORA	0.069		GENTE	0.038	0.004
GENTE	0.038	0.039	GRANDE		0.022
GRANDE		0.020	HUMO	0.125	0.041
HUMO	0.125	0.076	LATA	0.036	
INORGANICO		0.031	LIMPIA	0.072	0.009
LATA	0.036	0.041	MEDIO		0.014
LIMPIA	0.072	0.140	METROPOLI	0.030	
METROPOLI	0.030		NATURAL	0.316	0.432
NATURAL	0.316	0.383	ORGANICO	0.031	
ORGANICO	0.031		OXIGENO	0.087	0.067
OXIGENO	0.087	0.027	PAPEL	0.110	0.114
PAPEL	0.110	0.168	PERSONA	0.171	0.098
PERSONA	0.171	0.154	PLANTA	0.794	0.621
PLANTA	0.794	1.000	PLÁSTICO	0.152	0.121
PLÁSTICO	0.152	0.268	POBLACION		0.057
RECICLAJE	0.125	0.138	RECICLAJE	0.125	0.090
REUSAR	0.278	0.191	REUSAR	0.278	0.130
RIO	0.036		RIO	0.036	
RUIDO		0.074	RUIDO		0.017
SEMILLA	0.020		SEMILLA	0.020	
SMOG	0.063	0.037	SMOG	0.063	0.042
SUCIO	0.112	0.117	SUCIO	0.112	0.065
TALA		0.001	TALA		0.001
TIENDAS	0.001		TIENDAS	0.001	
TIERRA	0.041	0.264	TIERRA	0.041	0.138
TOXICO	0.001		TOXICO	0.001	
TRAFICO	0.081	0.015	TRAFICO	0.081	
VERDE	0.493	0.463	VERDE	0.493	0.205
VIDA	0.049	0.130	VIDA	0.049	0.208
VIDRIO		0.029			

La Tabla 4.2 expone los resultados del cálculo de similitud de todas la comparaciones entre las RSN de principiantes aplicando las medidas de similitud cosenoidal, coeficiente de Dice y coeficiente de Jaccard.

Tabla 4. 2. Resultado de similitud obtenido en la comparación entre las RSN de principiantes.

COSENOIDAL				
	eco1	eco2	eco3	eco4
eco1	1.000	0.942	0.930	0.927
eco2	0.942	1.000	0.913	0.954
eco3	0.930	0.913	1.000	0.939
eco4	0.927	0.954	0.939	1.000

DICE				
	eco1	eco2	eco3	eco4
eco1	1.000	0.940	0.928	0.919
eco2	0.940	1.000	0.913	0.936
eco3	0.928	0.913	1.000	0.925
eco4	0.919	0.936	0.925	1.000

JACCARD				
	eco1	eco2	eco3	eco4
eco1	1.000	0.886	0.867	0.851
eco2	0.886	1.000	0.839	0.879
eco3	0.867	0.839	1.000	0.861
eco4	0.851	0.879	0.861	1.000

4.4.2. Análisis de resultados de comparación entre RSN de principiantes

En el caso de la comparación entre principiantes el valor de similitud obtenido es muy parecido como se puede observar en la Tabla 4.2, los valores de similitud cosenoidal varían entre 0.913 y 0.954, el valor mínimo de similitud se obtuvo en la comparación entre la RSN ECOLOGÍA 3 y la RSN ECOLOGÍA 4 y el valor máximo de similitud se dio entre la RSN ECOLOGÍA 2 y la RSN ECOLOGÍA 4.

En la Tabla 4.3 se expone información con respecto a los valores de similitud y la proporción de definidores coincidentes y discrepantes para las comparaciones entre principiantes que obtuvieron la similitud máxima y mínima, como se puede observar, existe un valor muy pequeño de diferencia entre estos resultados (0.041), asimismo se puede identificar que la cantidad y porcentaje

de definidores coincidentes es menor en la comparación que obtuvo la mayor similitud (ECOLOGÍA 2-ECOLOGÍA 4), como se había mencionado, en el MEV, el peso que se da a los términos (en este caso el valor normalizado M de los definidores) tiene un impacto en el resultado de la similitud entre RSN, por lo que se puede inferir que las diferencias en los valores de similitud calculados están derivadas de las diferencias existentes en los valores normalizados de M.

Tabla 4. 3. Valores relacionados con la similitud entre principiantes.

Comparación	Similitud	Definidores coincidentes		Definidores no coincidentes	
		Cantidad	Porcentaje	Cantidad	Porcentaje
ECOLOGÍA 2- ECOLOGÍA 3	0.913	35	71.43%	14	28.57%
ECOLOGÍA 2- ECOLOGÍA 4	0.954	34	70.83%	14	29.17%

Para comprobar lo anterior se realizó la sumatoria de las diferencias entre los valores normalizados M entre las redes comparadas y se obtuvo que para la comparación entre ECOLOGÍA 2 y ECOLOGÍA 4 el valor calculado fue 2.693 y para ECOLOGÍA 2 y ECOLOGÍA 3 el valor fue 3.340, es decir, para la comparación con la mayor similitud la diferencia entre los valores M es menor.

En la Figura 4.20 se muestra la representación de las redes visuales de ECOLOGÍA 2, ECOLOGÍA 3 y ECOLOGÍA 4 en las que se distinguen con color naranja los definidores que son comunes para ambas redes; aunque la diferencia de similitud entre las redes es mínima (.041), en la representación visual se puede identificar que existe una diferencia significativa entre el peso del definidor AGUA en la RSN ECOLOGÍA 2 y la RSN ECOLOGÍA 3, además el término TIERRA que no es común con ECOLOGÍA 3 tiene un valor M normalizado mayor que los definidores que no son coincidentes entre ECOLOGÍA 2 y ECOLOGÍA 4.

Tabla 4. 4. Muestra la matriz de comparación entre los definidores de las plantillas del Experto 1 y del Experto 2.

	Experto 1	Experto 2		Experto 1	Experto 2
AGUA		0.259	INTENSIDAD	0.158	
AIRE		0.185	INTERACCION	0.842	0.074
ALUMINIO		0.001	LUZ	0.283	
APROVECHAMIENTO		0.222	MAR	0.105	
APROVECHAR	0.105		METROPOLI	0.421	
ARBOL		0.407	MEZCLA	0.105	
ARBOLAR	0.283		NATURALEZA	0.388	
ARMONIA	0.105	0.222	NICHO		0.259
BASURA	0.211	0.185	NINO		0.074
BELLEZA		0.185	NUCLEAR		0.222
BIODIVERSIDAD	0.789		NUCLEO	0.053	
BIOSFERA		0.222	NUTRIENTE		0.001
BOSQUE	0.316	0.407	OCEANO		0.111
CALLE		0.001	ORGANISMO	0.421	
CARRO	0.158	0.222	OXIGENO		0.298
CASA	0.421	0.185	PAISAJE		0.148
CEMENTO	0.283		PANORAMICO	0.105	
CICLO	0.528		PARQUE		0.148
COMBUSTIBLE		0.037	PLANETA		0.298
COMPOSTA	0.001		PLANTA	0.421	
CONCIENCIA		1.000	PLASTICO		0.037
CONSERVACION		0.222	POBLACION	0.474	0.407
CONTAGIO	0.053		PODAS	0.105	
CONTAMINACION		0.111	PODER	0.316	
CONTAMINANTES		0.333	POLITICA		0.001
CONTENEDORES		0.074	POTENCIA	0.211	
CREATIVIDAD		0.148	PROCESAR	0.474	
CUIDADO		0.298	PUDRICION	0.158	
DESCOMPOSICION	0.421		PUREZA		0.148
DESECHO		0.298	QUIMICOS	0.388	
DESERTO		0.111	RECUPERAR	0.388	
DESTRUCCION		0.593	RECURSOS	0.528	
DETERIORO	0.474		REDUCIR	0.211	
EDIFICIO		0.148	REGENERACION		0.148
EMPRESA		0.333	REHABILITAR	0.316	
ENCAUSAR	0.053		RELACION	0.158	0.001
ENFERMEDAD		0.222	RESTAURANTE		0.037
EOLICA		0.185	RETOMAR	0.158	
EQUILIBRIO		0.704	REUSAR	0.283	
ESENCIA	0.053		RIESGO		0.259
ESPACIO		0.259	RURAL		0.074
ESTRES	0.211		SEÑALIZACION	0.001	
ESTRUCTURA	0.283		SIERRA	0.211	
ESTUDIO		0.148	SISTEMA	0.211	
FABRICA		0.444	SMOG	0.283	
FAUNA		0.074	SOL	0.388	
FLORA		0.111	SOLAR		0.259
FLUJO	0.388		SUELO		0.298
FORTALEZA	0.528		SUSTENTABILIDAD	1.000	
FUERTE	0.001		TIERRA	0.579	
GAS		0.333	TOXICO		0.259
GENTE	0.388		TRANQUILIDAD	0.001	
HABITAT	0.105		TRANSFORMAR	0.421	
HIDROCARBURO		0.111	TRANSPORTE		0.148
HIDROELECTRICA		0.148	UNIVERSO	0.895	
HOJA	0.158		UTILIZACION		0.333
HUMANO		0.407	VARIEDAD	0.211	
HUMO	0.316		VEGETACION		0.222
IDENTIDAD	0.001		VERDE	0.842	0.037
IMPULSO	0.105		VIDA	0.474	
INDICES		0.001	VIENTO	0.053	
INFECCION	0.001				
INJERTO	0.053				

Tabla 4. 5. Muestra la matriz de comparación entre los definidores de la plantilla del Experto 2 y del Experto 4.

	Experto 2	Experto 4		Experto 2	Experto 4
ABONO		0.160	GASOLINERIA		0.040
AGUA	0.259	1.000	GEOTERMICA		0.040
AGUAS RESIDUALES		0.200	HIDROCARBURO	0.111	
AIRE	0.185		HIDROELECTRICA	0.148	0.320
ALCANTARILLAS		0.080	HOMBRE		0.680
ALIMENTO		0.160	HUERTO		0.080
ALUMINIO	0.001	0.080	HUMANO	0.407	
AMBULANTE		0.120	INDICES	0.001	
APROVECHAMIENTO	0.222	0.320	INDUSTRIA		0.320
ARBOL	0.407		INTERACCION	0.074	
AREA VERDE		0.360	INVERNADERO		0.280
ARMONIA	0.222		LLUVIA		0.200
AVENIDA		0.160	LUZ		0.200
AZUFRE		0.001	MEDIO AMBIENTE		0.200
BASURA	0.185	0.560	NEGOCIO		0.160
BELLEZA	0.185		NICHO	0.259	
BIODIVERSIDAD		0.360	NINO	0.074	
BIOSFERA	0.222		NUCLEAR	0.222	0.001
BOSQUE	0.407		NUTRIENTE	0.001	0.200
CALLE	0.001		OCEANO	0.111	
CARRO	0.222	0.680	OXIGENO	0.296	
CARTON		0.160	OZONO		0.320
CASA	0.185		PAISAJE	0.148	
CLIMA		0.280	PAPEL		0.120
COMBUSTIBLE	0.037		PARQUE	0.148	
COMPOSTA		0.120	PLANETA	0.296	
COMUNIDAD		0.040	PLANTA		0.160
CONCIENCIA	1.000		PLASTICO	0.037	
CONSERVACION	0.222		POBLACION	0.407	0.560
CONSTRUCCION		0.160	POLITICA	0.001	
CONTAMINACION	0.111	0.640	PRODUCTIVIDAD		0.120
CONTAMINANTES	0.333		PUREZA	0.148	
CONTENEDORES	0.074		RECURSOS		0.280
CREATIVIDAD	0.148		REGENERACION	0.148	
CUIDADO	0.296		RELACION	0.001	
DESARROLLO		0.240	RESTAURANTE	0.037	0.120
DESECHO	0.296	0.200	RIESGO	0.259	
DESIERTO	0.111		RURAL	0.074	0.240
DESTRUCCION	0.593		SEMAFORO		0.001
EDIFICIO	0.148	0.360	SENALAMIENTO		0.040
ELECTRODOMESTICO		0.080	SMOG		0.280
EMPRESA	0.333		SOL		0.240
ENFERMEDAD	0.222	0.040	SOLAR	0.259	0.360
EOLICA	0.185		SUELO	0.296	0.920
EQUILIBRIO	0.704		SUSTENTABILIDAD		0.360
EROSION		0.001	TOXICO	0.259	
ESPACIO	0.259		TRANSPORTE	0.148	0.200
ESPECIES		0.001	URBANIZACION		0.200
ESTUDIO	0.148		UTILIZACION	0.333	
FABRICA	0.444		VEGETACION	0.222	0.360
FAUNA	0.074		VERDE	0.037	
FIERRO		0.001	VIALIDAD		0.240
FLORA	0.111		VIDRIO		0.040
FORRAJE		0.040			
GAS	0.333	0.520			

En la Tabla 4.6 se puede observar el resultado de la comparación entre todos los expertos una vez que se realizó el cálculo de los coeficientes de similitud.

Tabla 4. 6. Muestra los coeficientes de similitud como resultados de la comparación entre las plantillas de expertos.

COSENOIDAL				
	exp1	exp2	exp3	exp4
exp1	1.000	0.089	0.246	0.204
exp2	0.089	1.000	0.273	0.309
exp3	0.246	0.273	1.000	0.236
exp4	0.204	0.309	0.236	1.000

DICE				
	exp1	exp2	exp3	exp4
exp1	1.000	0.084	0.242	0.201
exp2	0.084	1.000	0.242	0.306
exp3	0.242	0.242	1.000	0.221
exp4	0.201	0.306	0.221	1.000

JACCARD				
	exp1	exp2	exp3	exp4
exp1	1.000	0.044	0.124	0.112
exp2	0.044	1.000	0.138	0.181
exp3	0.124	0.138	1.000	0.124
exp4	0.112	0.181	0.124	1.000

4.4.4. Análisis de resultados de comparación entre RSN de expertos

Se puede observar que los valores de similitud cosenoidal generados para la comparación entre expertos son pequeños, van desde 0.089 hasta 0.309, lo anterior es debido a la discrepancia entre los definidores de las redes. Como se puede observar en la Tabla 4.6, el valor menor de similitud es 0.089 y está dado entre la red del Experto 1 y la red del Experto 2, por otra parte el valor mayor de similitud se encuentra entre la red del Experto 2 y la red del Experto 4, su valor es de 0.309.

La Tabla 4.4 muestra la matriz de comparación que corresponde al valor menor de similitud obtenida, en ella se puede observar que esta matriz tiene 124 definidores de los cuáles son coincidentes 9; asimismo en la Tabla 4.5 se muestra la matriz de comparación para el mayor valor de similitud obtenida en la que se puede notar que la cantidad total de definidores en la matriz es 108 de los cuales 20 son coincidentes.

Los resultados anteriores se muestran en la Tabla 5.2 en la que se puede percibir que la red con una mayor cantidad de coincidencias obtiene el mayor valor de similitud.

Tabla 4. 7. Análisis de comparación entre expertos.

Comparación	Similitud	Definidores coincidentes		Definidores no coincidentes	
		Cantidad	Porcentaje	Cantidad	Porcentaje
EXPERTO 1- EXPERTO 2	0.089	9	7.26%	115	92.74%
EXPERTO 2- EXPERTO 4	0.309	20	18.52%	88	81.48%

En la Figura 4.21 se muestran las representaciones visuales de las RSN de Experto 1, Experto 2 y Experto 4 en las que se ha identificado con color naranja los definidores comunes entre las redes, se puede observar que entre Experto 1 y Experto 2, los cuáles obtuvieron el valor mínimo de similitud (0.089), existe menor cantidad de definidores coincidentes y estos definidores corresponden a valores M de menor peso y además existen diferencias entre estos valores; entre Experto 2 y Experto 4 existe una mayor cantidad de definidores coincidentes pero existe diferencia en los valores M de estos lo que lleva a un valor de similitud obtenido de 0.309.

Aunque todos los expertos son del campo de estudio de ecología, sus especialidades son distintas, el valor mayor de similitud de la comparación se obtuvo entre el experto de ecología forestal con el experto en ecología acuática, aunque su valor es 0.309. El valor menor de similitud se obtuvo entre el experto de educación ambiental y el de ecología forestal.

Tabla 4. 8. Matriz de comparación de RED TOTAL de principiantes con la red del EXPERTO 1.

	RED TOTAL	EXP 1		RED TOTAL	EXP 1
AGUA	0.255		LIMPIA	0.098	
AIRE	0.152		MAR		0.118
ALUMINIO	0.060		METROPOLI		0.471
AMBIENTE	0.344		MEZCLA		0.118
ANIMAL	0.711		NATURAL	0.406	
APROVECHAR		0.118	NATURALEZA		0.412
ARBOL	1.000		ORGANISMO		0.471
ARBOLAR		0.294	OXIGENO	0.074	
ARMONIA		0.118	PANORAMICO		0.118
BASURA	0.539	0.235	PAPEL	0.131	
BIODIVERSIDAD		0.882	PERSONA	0.138	
BOSQUE	0.177	0.353	PLANTA	0.778	0.471
CALLE	0.010		PLASTICO	0.179	
CARRO	0.365	0.176	POBLACION		0.529
CARTON	0.026		PODAS		0.118
CASA	0.081	0.471	PROCESAR		0.529
CEMENTO		0.294	PUDRICION		0.176
CICLO		0.588	QUIMICOS		0.412
CIUDAD	0.022		RECICLAJE	0.128	
COMPOSTA		0.001	RECUPERAR		0.412
CONTAGIO		0.059	RECURSOS		0.588
CONTAMINAC...	0.329		REDUCIR		0.235
CUIDADO	0.127		REHABILITAR		0.353
DESCOMPOSICION		0.471	RELACIONES		0.176
DETERIORO		0.529	RETOMAR		0.176
ECOLOGIA	0.112		REUSAR	0.186	0.294
ECOSISTEMA	0.216		RUIDO	0.024	
EDIFICIO	0.234		SENALIZACION		0.001
ENCAUSAR		0.059	SIERRA		0.235
ENFERMEDAD	0.032		SISTEMA		0.235
ESENCIA		0.059	SMOG	0.062	0.294
ESTRES		0.235	SUCIO	0.097	
ESTRUCTURA		0.294	SUSTENTABILIDAD		0.529
FABRICA	0.053		TALA	0.001	
FLORA	0.029		TIERRA		0.647
FLUJO		0.412	TRAFICO	0.031	
FORTALEZA		0.059	TRANQUILIDAD		0.001
GENTE	0.048	0.412	TRANSFORMAR		0.471
GRANDE	0.024		UNIVERSO		1.000
HOJA		0.176	VARIEDAD		0.235
HUMO	0.115	0.353	VERDE	0.326	0.941
HABITAT		0.118	VIDA	0.147	0.529
IDENTIDAD		0.001	VIENTO		0.059
INFECCION		0.001			
INJERTO		0.059			
INTERACCION		0.941			
LATA	0.027				

Tabla 4. 9. Matriz de comparación de RED TOTAL de principiantes con la red del EXPERTO 3.

	RED TOTAL	EXP 3		RED TOTAL	EXP 3
ABIOTICO		0.214	LATA	0.027	
AGUA	0.255	0.001	LETRERO		0.214
AIRE	0.152	0.643	LIMPIA	0.098	
ALUMINIO	0.060		LUZ		0.214
AMBIENTE	0.344		METROPOLI		0.571
ANIMAL	0.711	0.071	MONTANAS		0.071
ARBOL	1.000	0.643	NATURAL	0.406	
AREA VERDE		0.143	OCEANO		0.001
ARMONIA		0.143	ORGANISMO		0.143
ARTIFICIAL		0.286	OXIGENO	0.074	
BALANCE		0.357	PAISAJE		0.429
BASURA	0.539	1.000	PAPEL	0.131	
BIOSFERA		0.500	PARQUE		0.214
BIOTICO		0.286	PARTICULAR		0.429
BOSQUE	0.177	0.429	PERSONA	0.138	
CALLE	0.010		PLANETA		0.500
CARRO	0.365		PLANTA	0.778	0.500
CARTON	0.028		PLASTICO	0.179	0.143
CASA	0.081	0.571	POBLACION		0.357
CIENCIA		0.643	PROBLEMÁTICA		0.143
CIUDAD	0.022		RECICLAJE	0.128	
CLIMA		0.357	RECUPERAR		0.643
COLOR		0.500	REEVALUAR		0.214
CONSUMISMO		0.071	REGENERACION		0.001
CONTAMINAC...	0.329		RELACION		0.643
CUIDADO	0.127		RESIDUO		0.571
DANO		0.001	REUSAR	0.186	
ECOLOGIA	0.112		REUTILIZAR		0.643
ECOSISTEMA	0.216	0.214	RUIDO	0.024	0.357
EDIFICIO	0.234	0.429	SISTEMA		0.143
ENERGIA		0.071	SMOG	0.062	0.571
ENFERMEDAD	0.032		SOCIAL		0.429
ESTUDIO		0.429	SUBPRODUCTO		0.357
FABRICA	0.053		SUCIO	0.097	
FAUNA		0.571	SUELO		0.071
FLORA	0.029	0.643	TALA	0.001	
GENTE	0.048	0.143	TRAFICO	0.031	0.500
GRANDE	0.024		TOXICO		0.071
GRIS		0.286	URBANO		0.714
HABITAT		0.643	UTILIDAD		0.500
HOMBRE		0.286	VEGETACION		0.643
HOMEOSTASIS		0.071	VERDE	0.326	0.357
HUMO	0.115		VIDA	0.147	0.643
INVESTIGACION		0.286			

En la Tabla 4.10 se puede observar que el valor menor de similitud entre la RSN total de participantes y los expertos ocurrió con el Experto 1 y el valor mayor de similitud ocurrió con el Experto 3.

Tabla 4. 10. Resultado de cálculo de similitud entre los expertos y la RED TOTAL de principiantes.

	RED TOTAL PRINCIPIANTES		
	Cosenoidal	Dice	Jaccard
exp 1	0.203	0.179	0.098
exp 2	0.257	0.237	0.134
exp 3	0.365	0.308	0.182
exp 4	0.288	0.248	0.141

4.4.6. Análisis de resultados de comparación entre Red Total de principiantes con RSN de expertos

Entre los valores de similitud calculados en esta comparación existe poca variación, el valor de similitud máximo de la Red Total de Principiantes se obtuvo con el Experto 3 (Legislación ambiental) y fue de 0.365 y la el valor mínimo de similitud se obtuvo con el Experto 1 (Educación ambiental) con un valor de 0.203.

En la matriz de comparación de la Tabla 4.8 se identifica que existen 11 definidores coincidentes (12.22%) entre el experto 1 y la Red Total de Principiantes, asimismo en la Tabla 4.9 se encuentra que existen 18 definidores coincidentes (20.69%) entre el Experto 3 y la Red Total de Principiantes lo que se refleja en los valores de similitud obtenidos mediante el MEV aplicado en la comparación de RSN.

En la Figura 4.22 se muestran las redes visuales que corresponden a las comparaciones que obtuvieron el menor y mayor valor de similitud, en cada una de ellas se ha marcado con un color naranja los definidores que son coincidentes para identificar de manera cualitativa la similitud entre las redes,

en esta figura se aprecia visualmente que existe una cantidad mayor de definidores coincidentes en la comparación de mayor similitud calculada.

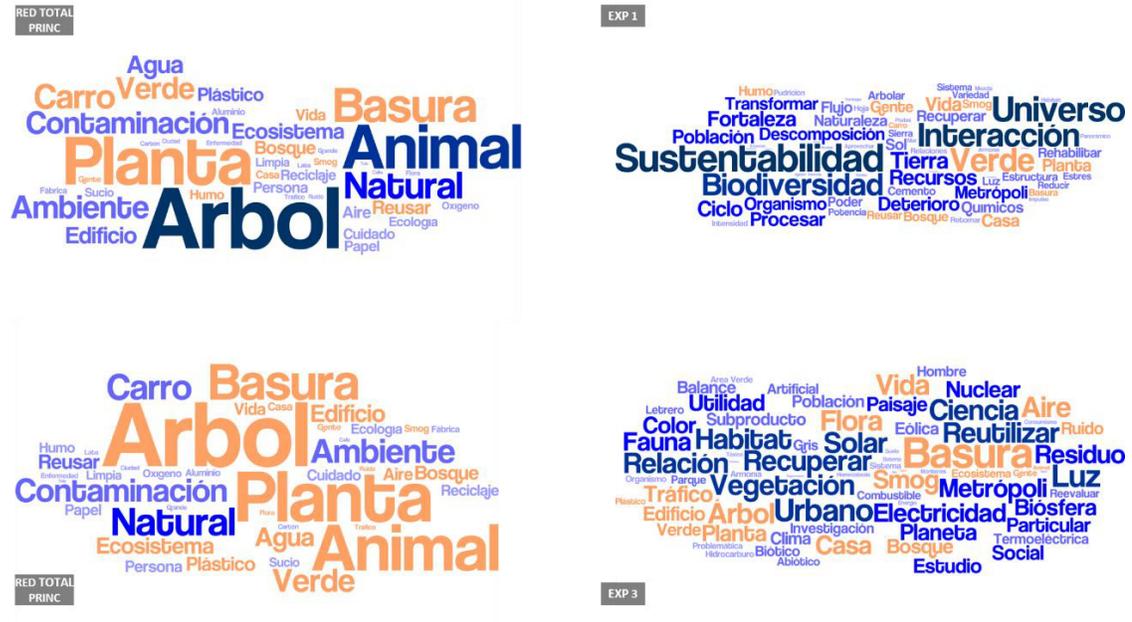


Figura 4. 22. Representación de redes visuales para comparación de similitud entre RED TOTAL de principiantes y expertos, en la figura se muestran las comparaciones que obtuvieron menor y mayor similitud.

La RED TOTAL de principiantes obtuvo una mayor similitud con el Experto 3 que tiene especialidad en legislación ambiental y obtuvo una menor similitud con el Experto 1 cuya especialidad es educación ambiental.

4.4.7. Comparación entre RSN de principiantes con Red Total de Expertos

Se realizó la comparación entre la RSN formada con los definidores de todos los expertos (Red Total de Expertos) con cada una de las RSN de principiantes para encontrar el valor de similitud entre ellas, las matrices de comparación generadas entre cada grupo de principiantes y la RSN de expertos se incluyen el Apéndice D.

En la Tabla 4.11 se muestran los resultados de los cálculos de similitud, en ella se puede observar que los valores de similitud varían entre 0.377, en la comparación de la Red Total de Expertos con ECOLOGÍA 4, y 0.448 con ECOLOGÍA 1, con base en la información anterior se conoce que la similitudes son muy parecidas ya que su variación es de .07, este parecido entre los valores de similitud se debe a que la similitud entre las RSN de los principiantes también tiene muy poca variación (.041) como se muestra en la subsección 4.4.1.

Tabla 4. 11. Valores de similitud calculados entre la RED TOTAL DE EXPERTOS y las RSN de principiantes.

	RED TOTAL EXPERTOS		
	Cosenoidal	Dice	Jaccard
eco1	0.448	0.447	0.288
eco2	0.378	0.377	0.233
eco3	0.403	0.403	0.252
eco4	0.377	0.369	0.227

En la Tabla 4.12 se muestra el porcentaje de definidores coincidentes en la comparación, se puede observar que los valores de similitud obtenidos son acordes con la cantidad de definidores coincidentes.

Tabla 4. 12. Información de resultados de comparación entre Red Total de Expertos y RSN de principiantes.

Comparación	Similitud	Definidores coincidentes		Definidores no coincidentes	
		Cantidad	Porcentaje	Cantidad	Porcentaje
RED TOTAL EXP- ECO 1	0.448	22	26.19%	62	73.81%
RED TOTAL EXP- ECO 2	0.378	19	23.46%	62	76.54%
RED TOTAL EXP- ECO 3	0.403	20	23.81%	64	76.19%
RED TOTAL EXP- ECO 4	0.377	19	22.89%	64	77.11%

La Figura 4.23 muestra la visualización de la Red Total de Expertos comparada con cada una de las redes de principiantes, en ella se puede percibir de manera visual, que es pequeña la variación de la similitud de estas comparaciones, ya que existe una gran similitud entre las redes de principiantes.

4.4.8. Comparación entre Red Total de Principiantes con Red Total de Expertos

Se realizó el proceso de comparación entre las redes semánticas naturales generadas con la totalidad de los principiantes y la totalidad de los expertos aplicando el MEV, en la Tabla 4.13 se muestra la matriz de comparación generada en la que se pueden identificar los definidores coincidentes entre ambas redes así como los valores M normalizados correspondientes de cada uno de ellos.

La Tabla 4.14 expone los resultados de similitud obtenidos entre la Red Total de Principiantes y La Red Total de Expertos y en la Figura 4.24 se muestran las redes visuales de la comparación en donde se observa la similitud cosenoidal calculada de 0.4.

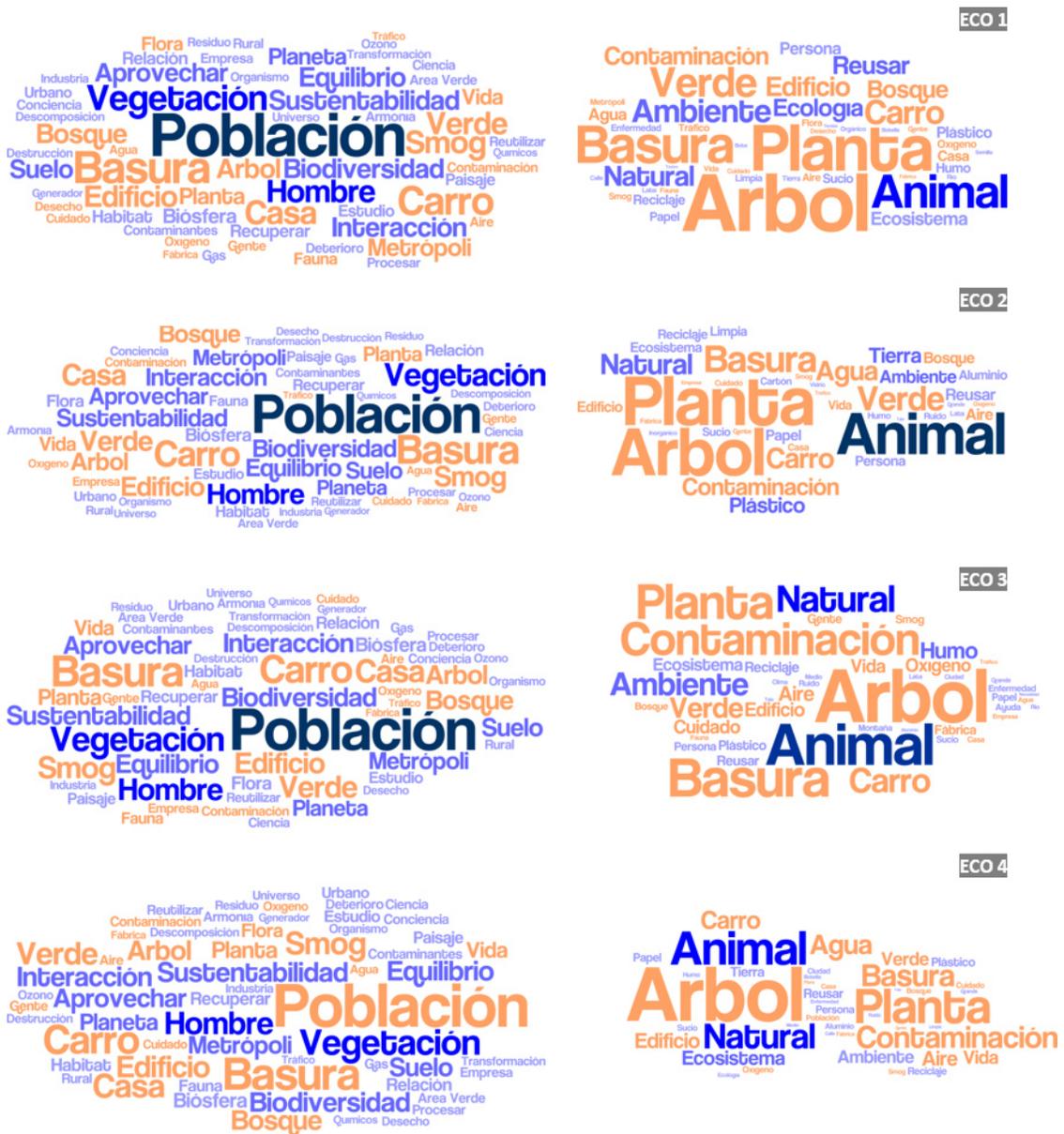


Figura 4. 23. Muestra del lado izquierdo la RED TOTAL DE EXPERTOS comparada con cada una de las redes de principiantes.

Tabla 4. 13. Matriz de comparación entre las redes semánticas naturales totales de expertos y principiantes.

	RED TOTAL PRINC	RED TOTAL EXP		RED TOTAL PRINC	RED TOTAL EXP
AGUA	0.255	0.028	HOMBRE		0.417
AIRE	0.152	0.056	HUMO	0.115	
ALUMINIO	0.060		INDUSTRIA		0.028
AMBIENTE	0.344		INTERACCION		0.333
ANIMAL	0.711		LATA	0.027	
APROVECHAR		0.306	LIMPIA	0.098	
ARBOL	1.000	0.333	METROPOLI		0.278
AREA VERDE		0.056	NATURAL	0.406	
ARMONIA		0.056	ORGANISMO		0.028
BASURA	0.539	0.694	OXIGENO	0.074	0.028
BIODIVERSIDAD		0.333	OZONO		0.028
BIOESFERA		0.194	PAISAJE		0.111
BOSQUE	0.177	0.333	PAPEL	0.131	
CALLE	0.010		PERSONA	0.138	
CARRO	0.365	0.583	PLANETA		0.250
CARTON	0.026		PLANTA	0.778	0.250
CASA	0.081	0.444	PLASTICO	0.179	
CIENCIA		0.056	POBLACION		1.000
CIUDAD	0.022		PROCESAR		0.056
CONCIENCIA		0.056	QUIMICOS		0.001
CONTAMINAC...	0.329	0.056	RECICLAJE	0.128	
CONTAMINANTES		0.056	RECUPERAR		0.139
CUIDADO	0.127	0.028	RELACION		0.139
DESCOMPOSICION		0.028	RESIDUO		0.028
DESECHO		0.028	REUSAR	0.186	
DESTRUCCION		0.028	REUTILIZAR		0.056
DETERIORO		0.056	RUIDO	0.024	
ECOLOGIA	0.112		RURAL		0.056
ECOSISTEMA	0.216		SMOG	0.062	0.417
EDIFICIO	0.234	0.389	SUCIO	0.097	
EMPRESA		0.056	SUELO		0.306
ENFERMEDAD	0.032		SUSTENTABILIDAD		0.333
EQUILIBRIO		0.333	TALA	0.001	
ESTUDIO		0.111	TRAFICO	0.031	0.001
FABRICA	0.053	0.001	TRANSFORMACION		0.028
FAUNA		0.111	UNIVERSO		0.028
FLORA	0.029	0.167	URBANO		0.083
GAS		0.056	VEGETACION		0.500
GENERADOR		0.001	VERDE	0.326	0.417
GENTE	0.048	0.056	VIDA	0.147	0.222
GRANDE	0.024				
HABITAT		0.139			

Tabla 4. 14. Similitud entre RED TOTAL de principiantes y RED TOTAL de expertos.

	RED TOTAL PRINCIPIANTES - RED TOTAL DE EXPERTOS
Cosenoidal	0.400
Dice	0.398
Jaccard	0.248



Figura 4. 24. Visualización de la comparación entre Red Total de Principiantes y Red Total de Expertos.

4.4.9. Resultados de la relación entre las medidas de similitud utilizadas

En este trabajo se calcularon tres medidas de similitud entre cada una de las comparaciones realizadas; para analizar la relación entre estas medidas, de manera inicial se realizaron gráficas para observar su comportamiento. En la Figura 4.25 se muestra la gráfica de los coeficientes de similitud calculados en la comparación entre la Red Total de Expertos y cada una de las redes de principiantes, los valores de estos coeficientes se expusieron en la Tabla 4.11; los valores de similitud calculados entre la Red Total de Principiantes y cada uno de los expertos que se muestran en la Tabla 4.10 fueron graficados, en la Figura 4.26 se puede apreciar la gráfica obtenida.

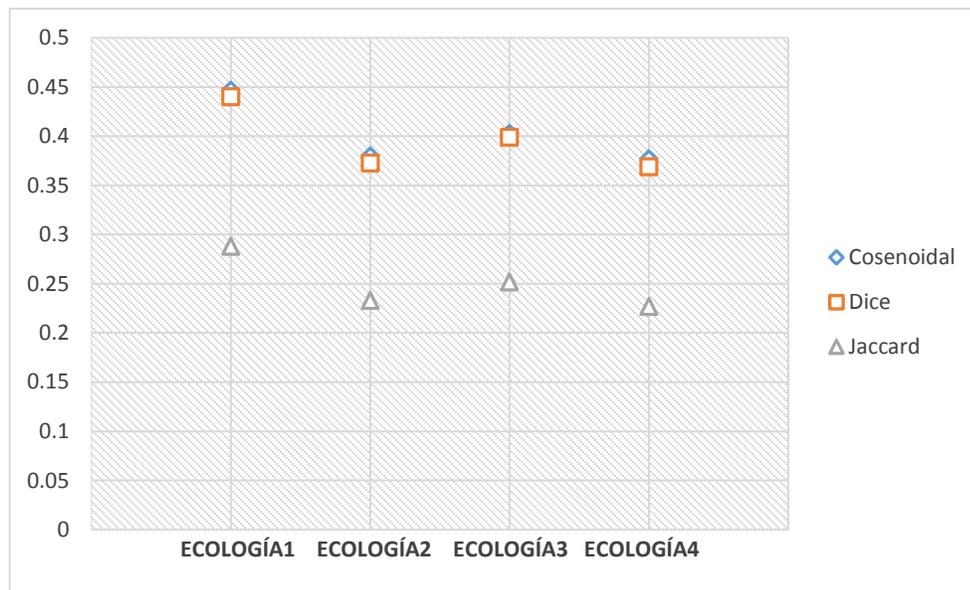


Figura 4. 25. Coeficientes de similitud calculados entre Red Total de Expertos y cada red de principiantes.

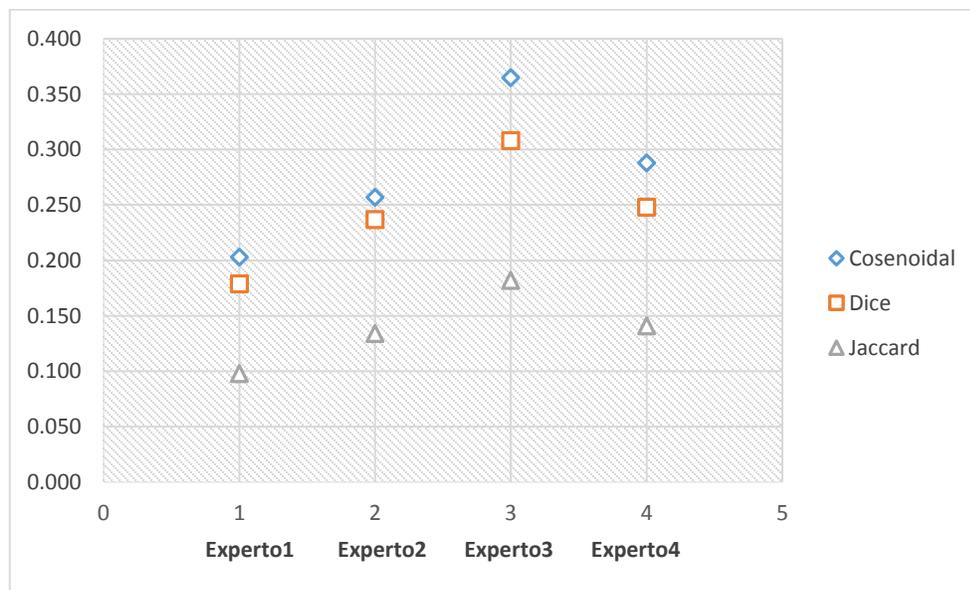


Figura 4. 26. Coeficientes de similitud calculados entre Red Total de Principiantes y cada uno de los expertos.

Como se puede observar existe una relación entre los coeficientes de las distintas medidas de similitud. Para medir y describir la relación entre dos variables se utiliza la correlación [86]. El coeficiente de correlación de Pearson genera valores entre -1 y 1 dado que es una covarianza estandarizada por la desviación estándar de las variables involucradas en la comparación [87]. Se calculó el coeficiente de correlación entre las medidas de similitud calculadas en todas las comparaciones de este trabajo (N=30), en la Tabla 4.12 se muestran los valores de los coeficientes de correlación calculados en donde se puede observar que se correlacionan significativamente con un nivel de significación establecido de 0.05.

Tabla 4. 15. Coeficientes de correlación de Pearson entre medidas de similitud.

	Cosenoidal	Dice	Jaccard
Cosenoidal	1	0.998	0.992
Dice	0.998	1	0.997
Jaccard	0.992	0.997	1

CAPÍTULO 5

CONCLUSIONES Y TRABAJO FUTURO

En esta investigación se propuso aplicar un Modelo de Recuperación de Información para realizar la comparación de Redes Semánticas Naturales. Con el objetivo de conseguir el material con el que se realizaría la investigación se generaron RSN con cuatro grupos de principiantes y cuatro expertos, estas redes fueron creadas mediante una aplicación web que almacena en una base de datos la información a manipular.

Con respecto a la hipótesis así como a las preguntas de investigación planteadas en este trabajo a continuación se expone las conclusiones obtenidas.

5.1. Conclusión del Modelo de Comparación de RSN

De los Modelos de Recuperación de Información revisados se seleccionó el Modelo del Espacio Vectorial para realizar la comparación entre RSN debido a que considera los pesos o importancia de los términos de consulta para hacer la búsqueda y no solamente la coincidencia de términos, y en las RSN se tiene el indicador denominado valor M el cual es interpretado como el peso del definidor.

De acuerdo a la definición del MEV, en este modelo se realiza la búsqueda de términos en documentos representados por vectores, lo cual se expresa con la ecuación 2.1

$$\vec{d}_j = (w_{1j}, w_{2j}, w_{3j}, w_{4j}, \dots, w_{nj})$$

Al considerar que una RSN se puede interpretar como un documento, cada definidor como un término del documento y el peso del término como el valor M del definidor, la ecuación 2.1 se transforma en:

$$\overrightarrow{rsn}_i = (M_{1i}, M_{2i}, M_{3i}, M_{4i}, M_{ni})$$

Fue posible adaptar el MEV trasladando el proceso que sigue este modelo hacia la comparación entre Redes Semánticas Naturales, las técnicas de comparación son las indicadas para MEV y el resultado de la similitud sería interpretado como el parecido que existe entre las redes comparadas, en la Figura 3.12 se representó la analogía para el proceso anterior.

Mediante un programa se realiza el pre-procesamiento de la información para obtener valores normalizados de M de cada RSN con los que se hace la comparación, se obtiene para cada comparación una matriz y el cálculo de tres distintas medidas de similitud: el coeficiente de cosenoidal, el coeficiente de Dice y el coeficiente de Jaccard.

Los resultados se analizaron con respecto a los valores calculados del coeficiente cosenoidal, aunque también se realizó un análisis del comportamiento de las otras medidas de similitud calculadas, este análisis se muestra en las secciones 4.4.9 y 5.2.

El análisis de resultados también se realizó con respecto a las coincidencias de la cantidad de definidores entre las redes y se encontró que eran acordes a los valores obtenidos del cálculo de la similitud cosenoidal en la mayoría de los casos, sólo se encontró en la comparación entre RSN de principiantes, cuyos valores son muy cercanos, que la mayor similitud no correspondía a la comparación con mayor cantidad de coincidencias como lo

muestra la Tabla 4.3 , lo anterior es posible que suceda ya que el método no sólo se basa en coincidencias, sino que también considera el Valor M normalizado para el cálculo.

Los resultados de la comparación entre las redes de principiantes son muy altos, sus valores van de 0.913 a 0.954, lo anterior es posible dado que las cuatro RSN fueron generadas con alumnos del mismo perfil, son alumnos de ingeniería de la misma institución y carrera, entre las redes hubo una coincidencia entre definidores también muy parecida que va del 70.83% al 71.43%.

En la comparación de similitud de su representación visual para el caso de comparación de mayor similitud (0.954) mostrada en la Figura 5.1 se puede identificar la gran similitud entre las RSN, destaca que los definidores discrepantes entre las redes son aquellos que tienen valores M pequeños, motivo por el cual, esta discrepancia no tuvo gran impacto en el valor de la similitud como se puede observar en la tabla incluida en la Figura 5.1.

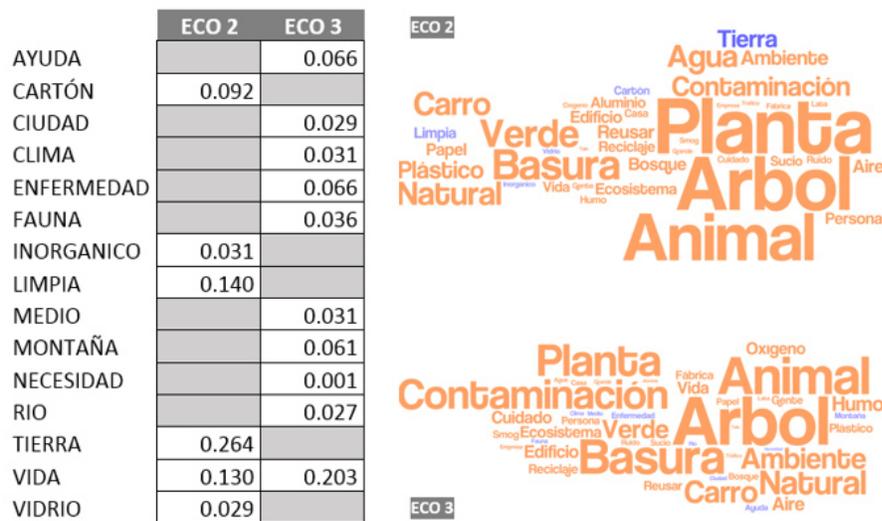


Figura 5. 1. Representación visual de comparación entre redes y tabla de definidores discrepantes.

Por lo anterior se considera que sí es posible comparar Redes Semánticas Naturales mediante el Modelo de Espacio Vectorial el cual es un modelo clásico de Recuperación de Información, esta comparación se puede realizar entre RSN de diferente cantidad de participantes ya que utiliza valores normalizados.

Se realizó la representación visual de las RSN mediante nubes de palabras, se pudo observar que mediante estas representaciones es posible identificar por su relevancia los definidores en la red.

Las comparaciones entre RSN también fueron representadas mediante las nubes de etiquetas, en donde se contrastaron los definidores comunes entre las redes, se observó que mediante este proceso es posible identificar de manera cualitativa la similitud entre las redes, se pudo observar que el valor de similitud está relacionado a la magnitud del valor M en los definidores de la red como se pudo observar en la Figura 4.20

5.2. Conclusión de las distintas medidas de similitud aplicadas en la comparación de RSN mediante el MEV

Los coeficientes de correlación que fueron calculados para determinar la relación entre las medidas de similitud los cuales son mostrados en la Tabla 4.12, indican que la correlación que existe entre las medidas de similitud es positiva muy fuerte, ya que el valor mínimo se obtuvo entre cosenoidal y Jaccard con una magnitud de 0.992 y el valor máximo se obtuvo entre cosenoidal y Dice con un valor de 0.998.

Con los resultados anteriores y la observación del comportamiento de las distintas medidas de similitud se puede sugerir que cualquiera de estas medidas de similitud puede ser utilizada en la comparación de RSN mediante el MEV ya que los resultados que se obtienen son coincidentes con respecto al orden posicional (*rank*); cuando se comparan RSN, es relevante conocer el orden posicional de la similitud entre ellas por lo que se puede considerar

indistinto utilizar cualquiera de las medidas siempre y cuando se utilice la misma medida en todas las comparaciones del mismo caso de estudio; sin embargo, es importante considerar que sí existe diferencia en la magnitud de similitud de las diferentes medidas motivo por el cual, es importante seleccionar una de las medidas para utilizarse en todas las comparaciones a realizar con la finalidad de que estos valores puedan ser comparados cuantitativamente.

Comentarios conclusivos

Se puede concluir que el Modelo de Espacio Vectorial adaptado a la comparación de RSN, es un método que se puede aplicar para realizar la comparación entre Redes Semánticas Naturales ya que se pudo comprobar que los resultados de similitud obtenidos reflejan el parecido que existe entre las RSN. La ventaja de este método es que puede ser aplicado entre redes de diferente tamaño ya que los valores son normalizados.

Con respecto a las medidas de similitud, de acuerdo con los resultados, se pueden aplicar cualquiera de ellas, ya que, aunque se profundizó en el análisis utilizando el cálculo de similitud cosenoidal, los resultados de similitud del coeficiente de Dice y del coeficiente de Jaccard tienen un comportamiento con una fuerte correlación con similitud cosenoidal por lo que proporcionarían resultados de comparación semejantes.

La generación de la visualización de la RSN mediante nube de etiquetas es una técnica útil para representar las redes, también permiten evaluar de manera cualitativa la similitud entre RSN. De las comparaciones de RSN realizadas en este trabajo, se seleccionaron aquellas que obtuvieron el valor de similitud más bajo (.089), el más alto (0.954) y un valor medio (0.463), con el objetivo de mostrar las comparaciones de las redes visuales de cada caso para observar de manera cualitativa su comparación, como se puede observar en la Figura 5.5.



Figura 5. 2. Comparación de visualización de RSN con distinta similitud.

La adaptación del Modelo de Espacio Vectorial de Recuperación de Información, para la comparación de Redes Semánticas Naturales tiene como resultado un método útil y comprensible para realizar la comparación entre RSN, aun cuando la cantidad de participantes sea distinta debido a la normalización de sus valores, lo anterior aunado a la generación de representaciones visuales mediante nubes de etiquetas, proporciona una nueva herramienta para la amplia variedad de campos en los que se usan la Redes Semánticas Naturales.

5.3. Trabajo Futuro

Como trabajo futuro se propone el estudio de otros Modelos de Recuperación de Información para comparar RSN con el objetivo de evaluar los resultados con respecto al Modelo de Espacio Vectorial aplicado en este trabajo; en esta investigación se utilizaron para la comparación de RSN medidas de similitud, sería de interés analizar medidas de distancia para revisar los resultados en la comparación. También podrían identificarse las métricas para cuestiones del estudio de distancia semántica progresiva.

Con respecto a la visualización de RSN se puede trabajar en el desarrollo de una aplicación que genere la red visual con nubes de etiquetas en la que se pueden realizar diferentes configuraciones para la comparación de la redes, también sería de interés realizar una representación visual en 3D para identificar las diferencias en la comparación de RSN.

REFERENCIAS

- [1] F. W. Lehmann y E. Y. Rodin, *Semantics Networks in Artificial Intelligence*. Vol 2, Pergamon Press, 1992.
- [2] J. F. Allen y A. M. Frisch, «What's in a semantic network?,» de 20th Annual meeting on Association for Computational Linguistics, Stroudsburg, PA, USA, 1982.
- [3] J. G. Figueroa, E. G. González y V. Solís, «Una aproximación al problema del significado: Las redes semánticas,» *Revista Latinoamericana de Psicología*, vol. 13, nº 3, pp:447-458, 1981.
- [4] A. González Lara y D. Navarro Reyes, «Evaluación de aprendizaje en un curso mediante técnica de Redes Semánticas Naturales,» *ANFEI Digital*, vol. 2, nº 3, pp. 1-10, 2015.
- [5] L. M. Casas García y R. Luengo González, «Redes Asociativas Pathfinder y Teoría de los Conceptos Nucleares. Aportaciones a la Investigación en Didáctica de las Matemáticas.,» de VII Simposio de la Sociedad Española de Investigación en Educación Matemática, Granada, 2003.
- [6] T. E. Goldsmith y K. Kraiger, «Applications of Structural Knowledge assessment to training evaluation,» de *Improving Training*

- Effectiveness In Organizations, New York, Psychology Press, 2014, pp. 73-95.
- [7] M. T. Chi, «Laboratory Methods for Assessing Experts' and Novices' Knowledge,» de The Cambridge Handbook of Expertise and Expert Performance, New York, Cambridge University Press, 2006, pp. 167-184.
- [8] T. Urwin, Information and Communication Technology for Development, Cambridge: Cambridge University Press, 2009.
- [9] R. Baeza-Yates y B. Ribeiro-Neto, Modern Information Retrieval, New York: ACM Press, 1999.
- [10] G. Salton, Automatic text processing: the transformation, analysis, and retrieval of information by computer, Boston: Addison-Wesley Longman Publishing Co., 1989.
- [11] U. Hanani, B. Shapira y P. Shoval, «Information filtering: Overview of Issues, Research and Systems,» User Modeling and User-Adapted Interaction, vol. 11, nº 3, pp. 203-259, 2001.
- [12] G. Salton y C. Buckley, «Term-weighting approaches in automatic text retrieval,» Information Processing & Management, vol. 24, nº 5, pp. 513-523, 1988.
- [13] C. Van Rijsbergen, Information Retrieval, Newton: Butterworth-Heinemann, 1979.
- [14] H. E. Arévalo Aguilera, «El uso de las redes semánticas naturales en las representaciones sociales de la responsabilidad,» Revista Internacional de Psicología, vol. 11, nº 2, 2010.

- [15] P. Murillo Estepa y S. Becerra Peña, «Las percepciones del clima escolar por directivos, docentes y alumnado mediante el empleo de redes semánticas naturales: su importancia en la gestión de los centros educativos,» *Revista de Educación*, vol. 350, pp. 375-379, 2009.
- [16] G. De la Cruz Flores y L. F. Abreu Hernández, «Atributos de tutores de posgrado por campo disciplinario. La perspectiva de estudiantes de la Universidad Nacional Autónoma de México,» *Perfiles educativos*, vol. 34, nº 138, pp. 10-27, 2012.
- [17] H. Trejo-Lucero, R. Camacho-Beiza y J. Herrera-Villalobos, «Significado semántico de «lactancia materna» y «lactancia artificial» en mujeres y hombres,» *Revista Mexicana de Pediatría*, vol. 78, nº 1, pp. 10-15, 2011.
- [18] E. Collipal, D. Cabalín, J. Vargas y H.Silva, «Conceptualización semántica del término anatomía humana por los estudiantes de medicina,» *International Journal of Morphology*, vol. 22, nº 3, pp. 185-188., 2004.
- [19] C. B. García y S. Jiménez Vidal, «Redes semánticas de los conceptos de presión y flotación en estudiantes de bachillerato,» *Revista Mexicana de Investigación Educativa.*, vol. 1, nº 2, pp. 343-361, 1996.
- [20] V. M. Padilla, M. G. Villarreal y M. C. Rodríguez, «Evaluación de indicadores de la representación estructural del conocimiento de un curso académico,» *Revista Mexicana de Psicología*, pp. 215-216, 2006.
- [21] V. M. Padilla y M. G. Villarreal, «Representación estructural del conocimiento entre estudiantes y maestros,» de XXXII Congreso

Nacional de Psicología: "Psicología: Una Ciencia Sin Fronteras", Mexicali, Baja California, 2005.

- [22] V. M. Padilla Montemayor, I. M. Sanders Von Arnim, M. C. Rodríguez Nieto, E. O. López Ramírez y M. G. Villarreal Peña, «Redes semánticas para medir el aprendizaje en línea y tradicional,» de Evaluación de redes semánticas, Argentina, EUDEM, 2009, pp. 157-182.
- [23] V. M. Padilla Montemayor y M. C. Rodríguez Nieto, «Evaluación de la representación del conocimiento de un curso,» de Congreso Latinoamericano Ciencias de la Educación, Cochabamba, 2011.
- [24] G. Hinojosa Rivero, «El tratamiento estadístico de las redes semánticas naturales,» Revista Internacional de Ciencias Sociales y Humanidades, vol. XVIII, nº 1, pp. 133-154, 2008.
- [25] M. G. Villarreal Peña, «La representación del aprendizaje en un curso y su evaluación estructural a través de la técnica de redes semánticas naturales,» de El enfoque cognitivo de la formación de nuestros significados, en prensa, pp. 58-73.
- [26] G. Salton y M. J. McGuill, Introduction to modern information retrieval, New York: Mc Graw Hill, 1986.
- [27] D. Hiemstra, «Information Retrieval Models,» de Information Retrieval: searching in the 21st Century, United Kingdom, Wiley, 2009, pp. 2-19.
- [28] H. Zhuge, «Interactive semantics,» Artificial Intelligence, vol. 174, nº 2, pp. 190-204, 2010.
- [29] N. Nilsson, Principles of Artificial Intelligence, Palo Alto: Tioga Press,

1980.

- [30] R. Brachman y H. Levesque, Knowledge Representation and Reasoning, San Francisco: Morgan Kaufmann, 2004.
- [31] E. A. Feigenbaum y J. Feldman, Computers & Thought, Cambridge, MA: MIT Press, 1995.
- [32] M. R. Quillian, Semantic Information Processing, Cambridge, Mass.: MIT Press, 1968.
- [33] M. Masterman, «Semantic message detection for machine translation, using an interlingua,» de International Conference on Machine Translation of Languages and Applied Language Analysis, Teddington, UK, 1961.
- [34] J. Sowa, «Semantic Networks,» de Encyclopedia of Cognitive Science, John Wiley & Sons, 2006.
- [35] D. Knuth, «Semantics of context-free languages,» Theory of Computing Systems, vol. 2, nº 2, p. 127–145, 1968.
- [36] G. Winskel, The Formal Semantics of Programming Languages: An Introduction, MIT Press, 1993.
- [37] E. Codd, «Extending the database relational model to capture more meaning,» ACM Transactions on Database Systems, vol. 4, pp. 397–434, 1979.
- [38] D. Harel y B. Rumpe, «Meaningful modeling: What’s the semantics of “semantics”?,» IEEE Computer, vol. 37, nº 10, pp. 64-72, 2004.
- [39] «UML® Resource Page,» Object Management Group, Inc, 2013. [En

- línea]. Available: <http://www.uml.org>. [Último acceso: 14 Julio 2014].
- [40] P. Chen, «The entity-relationship model, towards a unified view of data,» *ACM Transactions on Database Systems*, vol. 1, p. 9–36, 1976.
- [41] R. Breu, R. Grosu, F. Huber, B. Rumpe y W. Schwerin, «Towards a precise semantics for object-oriented modeling techniques,» de *European Conference on Object-Oriented Programming*, Heidelberg, 1998.
- [42] J. Rumbaugh, M. Blaha, W. Premerlani, F. Eddy y W. Lorensen, *Object-Oriented Modeling and Design*, Prentice Hall, 1991.
- [43] M. Klein, «XML, RDF, and relatives,» *Intelligent Systems*, vol. 16, nº 2, pp. 26-28, 2001.
- [44] J. G. Figueroa Nazuno, «Los orígenes conceptuales de la técnica de Redes Semánticas Naturales,» de *El enfoque cognitivo de la formación de nuestros significados*, Monterrey, México, Trillas, en prensa, pp. 15-29.
- [45] E. Vargas Medina, «Aspectos y fundamentos de la técnica de Redes Semánticas Naturales.,» de *El Enfoque Cognitivo de la formación de nuestros significados*, Monterrey, México, Trillas, en prensa, pp. 30-44.
- [46] J. L. Valdez M, *Las redes semánticas naturales, usos y aplicaciones en psicología social*, México: Universidad Autónoma del Estado de México, 1998.
- [47] J. Valdez y I. Reyes, «La Construcción de Instrumentos de Medición a partir de Categorías Semánticas,» *Revista de Psicología Social y*

Personalidad, vol. 1, nº 9, pp. 57-66, 1993.

- [48] I. Reyes-Lagunes, «Las Redes Semánticas Naturales: Su Conceptualización y su utilización en la construcción de instrumentos,» *Revista de Psicología Social y Personalidad*, vol. 9, nº 1, pp. 83-87, 1993.
- [49] A. Singhal, «Modern information retrieval: A brief overview.,» *IEEE Data Eng. Bull.*, vol. 24, nº 4, pp. 35-43, 2001.
- [50] V. Bush, «As We May Think,» *The Atlantic Monthly*, nº 176, pp. 101-108, 1945.
- [51] H. P. Luhn, «A statistical approach to mechanized encoding and searching of literary information,» *IBM Journal of Research and Development*, vol. 1, nº 4, pp. 309-317, 1957.
- [52] G. Salton y M. E. Lesk, «The SMART automatic document retrieval systems—an illustration,» *Communications of the ACM*, vol. 8, nº 6, pp. 391-398, 1965.
- [53] C. Cleverdon, «The Cranfield tests on index language devices,» *Aslib Proceedings*, vol. 19, nº 6, p. 173–192, 1967.
- [54] Association for Computing Machinery, «The ACM Computing Classification System,» 2016. [En línea]. Available: <http://dl.acm.org/ccs/ccs.cfm?id=10003317&lid=0.10002951.10003317&CFID=605310556&CFTOKEN=69526060>. [Último acceso: 12 Enero 2016].
- [55] G. H. Tolosa y F. R. A. Bordignon, *Introducción a la recuperación de información: conceptos, modelos y algoritmos básicos.*, Buenos Aires:

Universidad Nacional de Luján, 2008.

- [56] S. Dominich, «A unified mathematical definition of classical information retrieval,» *Journal of the American Society for Information Science*, vol. 51, nº 7, pp. 614-624, 2000.
- [57] C. D. Manning, P. Raghavan y H. Schütze, *An Introduction to Information Retrieval.*, Cambridge: Cambridge University Press, 2009.
- [58] R. Baeza-Yates y B. Ribeiro-Neto, *Modern Information Retrieval: The Concepts and Technology behind Search*, Harlow: Addison-Wesley, 2010.
- [59] G. Salton, A. Wong y C. S. Yang, «A Vector Space Model for Information Retrieval,» *Communications of the ACM*, vol. 18, nº 11, p. 613–620, 1975.
- [60] N. Fuhr, «Probabilistic Models in Information Retrieval,» *Computer Journal*, vol. 35, nº 3, pp. 243-255, 1992.
- [61] D. Hand, H. Mannila y P. Smyth, *Principles of Data Mining*, USA: MIT Press, 2001.
- [62] R. Duda, P. Hart y D. Stork, *Pattern Classification*, USA: Wiley, 2000.
- [63] P.-N. Tan, M. Steinbach y V. Kumar, *Introduction to Data Mining*, EUA: Addison Wesley, 2006.
- [64] C. Böhm, F. Krebs y H. Kriegel, «A Generic Technique for the Similarity Join,» de *Data Warehousing and Knowledge Discovery*, Dawak, 2002.
- [65] D. Massart y L. Kaufman, *Interpretation of Analytical Data by the Use*

of Cluster, New York: Wiley, 1983.

- [66] L. Muflikhah y B. Baharudin, «Document clustering using concept space and cosine similarity measurement,» de 6th International Conference on Computer Technology and Development, Hongkong, 2009.
- [67] D. A. Grossman y O. Frieder, Information retrieval: Algorithms and Heuristics, Amherst: Springer Science & Business Media, 2012.
- [68] C. D. Manning y H. Schütze, Foundations of Statistical Natural Language Processing, Boston: MIT Press, 2000.
- [69] S. H. Cha, «Comprehensive survey on distance/similarity measures between probability density functions,» International Journal Of Mathematical Models And Methods In Applied Sciences, vol. 1, nº 4, pp. 300-307, 2007.
- [70] J. D. Fekete, J. J. Van Wijk, J. T. Stasko y C. North, «The value of information visualization,» de Information Visualization , Heidelberg, Springer, 2008, pp. 1-18.
- [71] A. I. Zermeño, A. C. Arellano y V. A. Ramírez, «REDES SEMÁNTICAS NATURALES: técnica para representar los significados que los jóvenes tienen sobre televisión, Internet y expectativas de vida,» Estudios sobre las Culturas Contemporáneas, vol. XI, nº 22, pp. 305-334, 2005.
- [72] A. J. Cañas et al, «Concept maps: Integrating knowledge and information visualization,» de Knowledge and Information Visualization, Heidelberg, Springer, 2005, pp. 205-219.

- [73] B. Cabrero y S. Vidal, «Redes semánticas de los conceptos de presión y flotación en estudiantes de bachillerato,» *Revista Mexicana de Investigación Educativa*, vol. 1, nº 2, pp. 343-361, 1996.
- [74] R. Schaveneldt, *Pathfinder Associative Networks: Studies in Knowledge Organization*, New Jersey: Ablex, 1990.
- [75] Interlink, «The Knowledge Network Organizing Tool,» Interlink, 25 Abril 2015. [En línea]. Available: <http://interlinkinc.net/KNOT.html>. [Último acceso: 3 Enero 2016].
- [76] F. Torres Guerrero y E.O. López Ramírez, «Rastreo de la Información en Páginas Web a través del Significado,» *Daena: International Journal of Good Conscience*, vol. 5, nº 2, pp. 308-323, 2010.
- [77] J. A. Gil Pascual, *Técnicas e instrumentos para la recogida de información*, Madrid: UNED, 2011.
- [78] Ingeniería Cognitiva, «RSNweb,» FIME UANL, 2 Agosto 2014. [En línea]. Available: <http://ingenieriacognitiva.com/developer/redes/registro/>. [Último acceso: 6 Marzo 2015].
- [79] L. Shalabi, Z. Shaaban y B. Kasasbeh, «Data Mining: A Preprocessing Engine,» *Journal of Computer Science*, vol. 2, nº 9, pp. 735-739, 2006.
- [80] V. Cloquell, M. Santamarina y A. Hospitaler, «Nuevo procedimiento para la normalización de valores numéricos en la toma de decisiones,» de XVII Congreso Nacional de Ingeniería de Proyectos, Murcia, 2001.
- [81] N. Karthikeyani Visalakshi y K. Thangavel, «Impact of Normalization in Distributed K-Means Clustering,» *International Journal of Soft*

Computing, vol. 4, n° 4, pp. 168-172, 2009.

- [82] S. Bateman, C. Gutwin y M. Nacenta, «Seeing things in the clouds: the effect of visual features on tag cloud selections,» de nineteenth ACM conference on Hypertext and hypermedia, Pittsburgh, 2008.
- [83] J. Feinberg, «Wordle,» 2014. [En línea]. Available: <http://www.wordle.net/>. [Último acceso: 2 Enero 2016].
- [84] S. Dominich, Mathematical Foundations of Information Retrieval, Dordrecht: Kluwer Academic Publishers, 2001.
- [85] D. Alimohammadi, Mathematics for Classical Information Retrieval: Roots and Applications, Lincoln: University of Nebraska–Lincoln Libraries, 2010.
- [86] F. J. Gravetter y L. B. Wallnau, Statistics for The Behavioral Sciences, Boston: Cengage Learning, 2016.
- [87] M. Rodríguez Salazar y S. Alvarez Hernández, Coeficientes de asociación UAM 1, México: UAM 1, 2001.

ÍNDICE DE FIGURAS

1. 1. Representa una Red Semántica Natural generada para la temática de Simulación de Sistemas	2
1. 2. Proceso de Recuperación de Información.	4
2. 1. Árbol de Porfirio, dibujado por Peter de España (1329).....	11
2. 2. Tópicos del área de estudio de Recuperación de Información.	15
2. 3. Operaciones realizadas en Sistema de Recuperación de Información.....	16
2. 4. El Modelo de Espacio Vectorial representa documentos y consultas como vectores para compararlos.	20
2. 5. Visualización de RSN mediante mapa conceptual.....	27
2. 6. Visualización de la red semántica natural para representación del clima escolar	27
2. 7. Representación visual de RSN del tema de ecología generada mediante la aplicación KNOT.....	28
3. 1. Muestra la pantalla de inicio de RSNweb para el registro de información de los participantes en el estudio.....	36
3. 2. Pantalla de Instrucciones que muestra RSNweb al participante.....	37
3. 3. Interfaz de la aplicación	38

3. 4. Interfaz de RSNweb para limpieza de definidores de la red.	39
3. 5. Grupos SAM de RSN del Proyecto ECOLOGÍA 1	40
3. 6. Muestra la Plantilla que representa la RSN del experto 1	41
3. 7. Formación de vectores inicial e inicialM con los definidores y los valores M de la RSN para pre-procesamiento de la información.	43
3. 8. Muestra la salida del programa de pre-procesamiento para obtener los vectores para la comparación.	45
3. 9. Matices del color azul seleccionados para la representación visual.	46
3.10. Algoritmo para asignación de matiz de color de acuerdo al valor M para la visualización en nubes de etiquetas de RSN.....	46
3.11. Representación visual de RSN mediante nube de etiquetas.	48
3.12. Representa la analogía entre el proceso del Modelo de Espacio Vectorial clásico de Recuperación de Información y el proceso del Modelo de Espacio vectorial aplicado a la comparación de RSN.	49
3.13. En esta matriz cada fila representa una RSN, cada columna representa un definidor y cada elemento representa el Valor M de cada definidor en la RSN correspondiente.	50
3.14. Representa el cálculo de la similitud cosenoidal, y los coeficientes de Dice y Jaccard para la comparación de las RSN.....	51
4. 1. Representación de grupos SAM de RSN del proyecto ECOLOGÍA 1.....	53
4. 2. Representación de grupos SAM de RSN del proyecto ECOLOGÍA 2.....	54
4. 3. Representación de grupos SAM de RSN del proyecto ECOLOGÍA 3.....	54
4. 4. Representación de grupos SAM de RSN del proyecto ECOLOGÍA 4.....	55

4. 5. Representación de la RED TOTAL formada con la información obtenida de todos los principiantes.	56
4. 6. Muestra la Plantilla que representa la RSN del Experto 1.	57
4. 7. Muestra la Plantilla que representa la RSN del Experto 2.	57
4. 8. Muestra la Plantilla que representa la RSN del Experto 3.	58
4. 9. Muestra la Plantilla que representa la RSN del Experto 4.	58
4. 10. Representación de Red Total de Expertos generada con la información de los cuatro expertos participantes.	59
4. 11. Representación visual de RSN ECOLOGÍA 1 con nube de etiquetas	60
4. 12. Representación visual de RSN ECOLOGÍA 2 con nube de etiquetas. ...	60
4. 13. Representación visual de RSN ECOLOGÍA 3 con nube de etiquetas. ...	61
4. 14. Representación visual de RSN ECOLOGÍA 4 con nube de etiquetas. ...	61
4. 15. Representación visual mediante nubes de etiquetas de la plantilla de Experto 1.	62
4. 16. Representación visual mediante nubes de etiquetas de la plantilla de Experto 2.	63
4. 17. Representación visual mediante nubes de etiquetas de la plantilla de Experto 3.	63
4. 18. Representación visual mediante nubes de etiquetas de la plantilla de Experto 4.	64
4. 19. Visualización de Red Total de Principiantes y de Red Total de Expertos mediante nube de etiquetas.	64
4. 20. Representación de redes visuales de comparaciones de principiantes con menor y mayor similitud para identificación cualitativa de similitud. ...	69

4. 21. Redes visuales de Expertos para observar comparación entre la menor similitud y la mayor similitud obtenida.	74
4. 22. Representación de redes visuales para comparación de similitud entre RED TOTAL de principiantes y experto.....	78
4. 23. Muestra del lado izquierdo la RED TOTAL DE EXPERTOS comparada con cada una de las redes de principiantes.	81
4. 24. Visualización de la comparación entre Red Total de Principiantes y Red Total de Expertos.	83
4. 25. Coeficientes de similitud calculados entre Red Total de Expertos y cada red de principiantes.	84
4. 26. Coeficientes de similitud calculados entre Red Total de Principiantes y cada uno de los expertos.	84
5. 1. Representación visual de comparación entre redes y tabla de definidores discrepantes.	88
5. 2. Comparación de visualización de RSN con distinta similitud.	91

ÍNDICE DE TABLAS

2. 1. Muestra un ejemplo de los definidores jerarquizados para el concepto estímulo familia mediante la técnica RSN.	14
2. 2. Familia Minkowski de medidas de disimilitud.....	26
3.1. Representa la cantidad de participantes por cada RSN generada.	30
3.2. Matriz P que almacena la frecuencia acumulada definidor/jerarquía y cálculo de Valor M para 10 participantes.....	32
3. 3. Muestra la selección del Grupo SAM y el Valor FMG de cada definidor en el Grupo SAM.....	34
3. 4. Muestra el área de especialidad de los expertos en Ecología que contribuyeron al estudio.	40
3. 5. Información requerida para generar la visualización de la RSN ECOLOGÍA 1 mediante wordle.....	47
3. 6. Ejemplo de matriz de salida para la comparación de RSN mediante MEV.....	50
4. 1. Muestra la matriz de comparación entre la RSN ECOLOGÍA 1 con la RSN ECOLOGÍA 2 y de la RSN ECOLOGÍA1 con la RSN ECOLOGÍA 4.	66

4. 2. Resultado de similitud obtenido en la comparación entre las RSN de principiantes.	67
4. 3. Valores relacionados con la similitud entre principiantes.	68
4. 4. Muestra la matriz de comparación entre los definidores de las plantillas del Experto 1 y del Experto 2.	70
4. 5. Muestra la matriz de comparación entre los definidores de la plantilla del Experto 2 y del Experto 4.	71
4. 6. Muestra los coeficientes de similitud como resultados de la comparación entre las plantillas de expertos.	72
4. 7. Análisis de comparación entre expertos.	73
4. 8. Matriz de comparación de RED TOTAL de principiantes con la red del EXPERTO 1.	75
4. 9. Matriz de comparación de RED TOTAL de principiantes con la red del EXPERTO 3.	76
4. 10. Resultado de cálculo de similitud entre los expertos y la RED TOTAL de principiantes.	77
4. 11. Valores de similitud calculados entre la RED TOTAL DE EXPERTOS y las RSN de principiantes.	79
4. 12. Información de resultados de comparación entre Red Total de Expertos y RSN de principiantes.	79
4. 13. Matriz de comparación entre las redes semánticas naturales totales de expertos y principiantes.	82
4. 14. Similitud entre RED TOTAL de principiantes y RED TOTAL de expertos.	82
4. 15. Coeficientes de correlación de Pearson entre medidas de similitud.	85

A. 1. Formación de vectores inicial e inicialM para ECOLOGÍA 2.....	111
A. 2. Formación de vectores inicial e inicialM para ECOLOGÍA 3.....	111
A. 3. Formación de vectores inicial e inicialM para ECOLOGÍA4.....	112
A. 4. Formación de vectores inicial e inicialM para Red Total de Principiantes.....	112
B. 1. Vectores de salida del pre-procesamiento para ECOLOGÍA 2.....	113
B. 2. Vectores de salida del pre-procesamiento para ECOLOGÍA 3.....	113
B. 3. Vectores de salida del pre-procesamiento para ECOLOGÍA 4.....	114
B. 3. Vectores de salida del pre-procesamiento para Red Total de Principiantes.....	114
C. 1. Matriz de comparación entre ECOLOGÍA 1 y ECOLOGÍA 3.....	115
C. 2. Matriz de comparación entre ECOLOGÍA 2 y ECOLOGÍA 3.....	115
C. 3. Matriz de comparación entre ECOLOGÍA 2 y ECOLOGÍA 4.....	116
C. 4. Matriz de comparación entre ECOLOGÍA 3 y ECOLOGÍA 4.....	116
D. 1. Matriz de Comparación entre ECOLOGÍA 1 y la Red Total de Expertos.....	117
D. 2. Matriz de Comparación entre ECOLOGÍA 2 y la Red Total de Expertos.....	118
D. 3. Matriz de Comparación entre ECOLOGÍA 3 y la Red Total de Expertos.....	119
D. 4. Matriz de Comparación entre ECOLOGÍA 4 y la Red Total de Expertos.....	120

APÉNDICE A

Entrada para pre-procesamiento de información

Vector de definidores inicial[70] y vector de valor M por definidor inicialMI[70] para entrada al programa de pre-procesamiento de información de RSN generadas.

Tabla A. 1. Formación de vectores inicial e inicialM para ECOLOGÍA 2.

inicial[70]	inicialM[70]		
AGUA	59	HUMO	126
AGUA	75	INORGANICO	86
AGUA	105	LATA	95
AGUA	137	LIMPIA	75
AIRE	90	LIMPIA	108
AIRE	129	NATURAL	201
ALUMINIO	182	NATURAL	134
AMBIENTE	166	NATURAL	62
AMBIENTE	81	OXIGENO	83
ANIMAL	135	PAPEL	207
ANIMAL	210	PERSONA	195
ANIMAL	351	PLANTA	214
ANIMAL	105	PLANTA	155
ARBOL	102	PLANTA	301
ARBOL	176	PLANTA	272
ARBOL	237	PLASTICO	296
ARBOL	396	RECICLAJE	181
BASURA	246	REUSAR	228
BASURA	270	RUIDO	71
BOSQUE	99	RUIDO	53
BOSQUE	114	SMOG	92
CARRO	247	SUCIO	162
CARRO	143	TALA	59
CARTÓN	140	TIERRA	66
CASA	133	TIERRA	82
CONTAMINACIÓN	230	TIERRA	83
CONTAMINACIÓN	99	TIERRA	61
CUIDADO	108	TRAFICO	72
ECOSISTEMA	114	VERDE	139
ECOSISTEMA	77	VERDE	87
EDIFICIO	204	VERDE	156
EMPRESA	72	VERDE	86
FABRICA	93	VIDA	102
GENTE	93	VIDA	72
GRANDE	77	VIDRIO	85

Tabla A. 4. Formación de vectores inicial e inicialM para ECOLOGÍA 3.

inicial[70]	inicialM[70]		
AGUA	53	ENFERMEDAD	72
AIRE	67	FABRICA	105
AIRE	58	FAUNA	55
AIRE	40	GENTE	99
ALUMINIO	37	GRANDE	49
AMBIENTE	124	HUMO	180
AMBIENTE	84	LATA	51
AMBIENTE	48	MEDIO	52
ANIMAL	80	MONTAÑA	69
ANIMAL	153	NATURAL	151
ANIMAL	226	NATURAL	93
ARBOL	106	NATURAL	50
ARBOL	73	NECESIDAD	35
ARBOL	140	OXIGENO	47
ARBOL	273	OXIGENO	97
AYUDA	72	PAPEL	83
BASURA	255	PERSONA	94
BASURA	172	PLANTAS	54
BOSQUES	80	PLANTAS	57
CARRO	161	PLANTAS	165
CARRO	108	PLANTAS	113
CASA	62	PLASTICO	108
CIUDAD	51	RECICLAJE	114
CLIMA	52	REUSAR	111
CONTAMINACIÓN	172	RIO	50
CONTAMINACIÓN	48	RUIDO	66
CONTAMINACIÓN	99	SMOG	90
CONTAMINACIÓN	42	SUCIO	73
CUIDADO	55	TALA	39
CUIDADO	50	TRAFICO	47
CUIDADO	44	VERDE	116
ECOSISTEMA	73	VERDE	78
ECOSISTEMA	65	VERDE	34
EDIFICIO	155	VIDA	86
EMPRESA	52	VIDA	62

Tabla A. 6. Formación de vectores inicial e inicialM para ECOLOGÍA 4

inicial[70]	inicialM[70]		
AGUA	81	EDIFICIO	288
AGUA	121	ENFERMEDAD	73
AGUA	171	FABRICA	78
AIRE	91	FLORA	65
AIRE	103	GENTE	55
AIRE	88	GRANDE	73
ALUMINIO	107	HUMO	91
AMBIENTE	104	LIMPIA	60
AMBIENTE	85	MEDIO	65
AMBIENTE	54	NATURAL	224
ANIMAL	84	NATURAL	160
ANIMAL	191	NATURAL	93
ANIMAL	301	OXIGENO	117
ANIMAL	64	PAPEL	163
ARBOL	162	PERSONA	148
ARBOL	183	PLANTAS	107
ARBOL	280	PLANTAS	106
ARBOL	411	PLANTAS	192
BASURA	268	PLANTAS	258
BASURA	156	PLASTICO	170
BOSQUE	102	POBLACION	107
BOTELLA	75	RECICLAJE	140
CALLE	65	REUSAR	179
CARRO	249	RUIDO	68
CARRO	90	SMOG	92
CASA	93	SUCIO	115
CIUDAD	104	TALA	51
CONTAMINACIÓN	233	TIERRA	90
CONTAMINACIÓN	119	TIERRA	97
CONTAMINACION	66	VERDE	90
CUIDADO	109	VERDE	104
ECOLOGÍA	81	VERDE	59
ECOSISTEMA	92	VIDA	72
ECOSISTEMA	80	VIDA	111
ECOSISTEMA	100	VIDA	73

Tabla A. 4. Formación de vectores inicial e inicialM para Red Total de Principiantes.

inicial[70]	inicialM[70]		
AGUA	198	ECOSISTEMA	363
AGUA	300	ECOSISTEMA	261
AGUA	429	EDIFICIO	865
AIRE	279	ENFERMEDAD	254
AIRE	339	FABRICA	318
ALUMINIO	339	FLORA	246
AMBIENTE	480	GENTE	303
AMBIENTE	249	GRANDE	229
AMBIENTE	266	HUMO	506
AMBIENTE	203	LATA	241
ANIMAL	378	LIMPIA	241
ANIMAL	650	LIMPIA	212
ANIMAL	1077	NATURAL	668
ANIMAL	201	NATURAL	520
ARBOL	487	NATURAL	198
ARBOL	549	OXIGENO	383
ARBOL	793	PAPEL	553
ARBOL	1350	PERSONA	574
BASURA	1012	PLANTA	477
BASURA	774	PLANTA	410
BOSQUE	271	PLANTA	808
BOSQUE	422	PLANTA	814
CALLE	188	PLASTICO	699
CARRO	855	RECICLAJE	544
CARRO	407	REUSAR	720
CARTÓN	238	RUIDO	229
CASA	403	SMOG	345
CIUDAD	225	SUCIO	451
CONTAMINAC...	777	TALA	158
CONTAMINAC...	374	TRAFICO	253
CUIDADO	323	VERDE	464
CUIDADO	219	VERDE	453
ECOLOGIA	246	VERDE	226
ECOLOGIA	251	VIDA	362
ECOSISTEMA	187	VIDA	240

APÉNDICE B

Salida de pre-procesamiento de información

Vectores de salida del programa de pre-procesamiento de información para cada uno de los proyectos.

Tabla B. 1. Vectores de salida del pre-procesamiento para ECOLOGÍA 2

final[J]	finalM[J]	finalMNor[J]
AGUA	376	0.359003
AIRE	219	0.181200
ALUMINIO	182	0.139298
AMBIENTE	247	0.212911
ANIMAL	801	0.840317
ARBOL	911	0.964892
BASURA	516	0.517554
BOSQUE	213	0.174405
CARRO	390	0.374858
CARTÓN	140	0.091733
CASA	133	0.083805
CONTAMINAC...	329	0.305776
CUIDADO	108	0.055493
ECOSISTEMA	191	0.149490
EDIFICIO	204	0.164213
EMPRESA	72	0.014723
FABRICA	93	0.038505
GENTE	93	0.038505
GRANDE	77	0.020385
HUMO	126	0.075878
INORGANICO	86	0.030578
LATA	95	0.040770
LIMPIA	183	0.140430
NATURAL	397	0.382786
OXIGENO	83	0.027180
PAPEL	207	0.167610
PERSONA	195	0.154020
PLANTA	942	1.000000
PLÁSTICO	296	0.268403
RECICLAJE	181	0.138165
REUSAR	228	0.191393
RUIDO	124	0.073613
SMOG	92	0.037373
SUCIO	162	0.116648
TALA	59	0.000000
TIERRA	292	0.263873
TRAFICO	72	0.014723
VERDE	468	0.463194
VIDA	174	0.130238
VIDRIO	85	0.029445

Tabla B. 21. Vectores de salida del pre-procesamiento para ECOLOGÍA 3.

final[J]	finalM[J]	finalMNor[J]
AGUA	53	0.032316
AIRE	165	0.233393
ALUMINIO	37	0.003591
AMBIENTE	256	0.396768
ANIMAL	459	0.761221
ARBOL	592	1.000000
AYUDA	72	0.066427
BASURA	427	0.703770
BOSQUES	80	0.080790
CARRO	269	0.420108
CASA	62	0.048474
CIUDAD	51	0.028725
CLIMA	52	0.030521
CONTAMINAC...	361	0.585278
CUIDADO	149	0.204668
ECOSISTEMA	138	0.184919
EDIFICIO	155	0.215440
EMPRESA	52	0.030521
ENFERMEDAD	72	0.066427
FABRICA	105	0.125673
FAUNA	55	0.035907
GENTE	99	0.114901
GRANDE	49	0.025135
HUMO	180	0.260323
LATA	51	0.028725
MEDIO	52	0.030521
MONTANA	69	0.061041
NATURAL	294	0.464991
NECESIDAD	35	0.000000
OXIGENO	144	0.195691
PAPEL	83	0.086176
PERSONA	94	0.105925
PLANTAS	389	0.635548
PLÁSTICO	108	0.131059
RECICLAJE	114	0.141831
REUSAR	111	0.136445
RIO	50	0.026930
RUIDO	66	0.055655
SMOG	90	0.098743
SUCIO	73	0.068223
TALA	39	0.007181
TRAFICO	47	0.021544
VERDE	228	0.346499
VIDA	148	0.202873

Tabla B. 29. Vectores de salida del pre-procesamiento para ECOLOGÍA 4.

final[J]	finalM[J]	finalMNor[J]
AGUA	373	0.326904
AIRE	282	0.234518
ALUMINIO	107	0.056853
AMBIENTE	243	0.194924
ANIMAL	640	0.597970
ARBOL	1036	1.000000
BASURA	424	0.378680
BOSQUE	102	0.051777
BOTELLA	75	0.024365
CALLE	65	0.014213
CARRO	339	0.292386
CASA	93	0.042640
CIUDAD	104	0.053807
CONTAMINAC...	418	0.372589
CUIDADO	109	0.058883
ECOLOGIA	81	0.030457
ECOSISTEMA	272	0.224365
EDIFICIO	288	0.240609
ENFERMEDAD	73	0.022335
FABRICA	78	0.027411
FLORA	65	0.014213
GENTE	55	0.004061
GRANDE	73	0.022335
HUMO	91	0.040609
LIMPIA	60	0.009137
MEDIO	65	0.014213
NATURAL	477	0.432487
OXIGENO	117	0.067005
PAPEL	163	0.113706
PERSONA	148	0.098477
PLANTAS	663	0.621320
PLASTICO	170	0.120812
POBLACION	107	0.056853
RECICLAJE	140	0.090355
REUSAR	179	0.129949
RUIDO	68	0.017259
SMOG	92	0.041624
SUCIO	115	0.064975
TALA	51	0.000000
TIERRA	187	0.138071
VERDE	253	0.205076
VIDA	256	0.208122

Tabla B. 4. Vectores de salida del pre-procesamiento para Red Total de Principiantes

final[J]	finalM[J]	finalMNor[J]
AGUA	927	0.254551
AIRE	618	0.152267
ALUMINIO	339	0.059914
AMBIENTE	1198	0.344257
ANIMAL	2306	0.711023
ARBOL	3179	1.000000
BASURA	1786	0.538894
BOSQUE	693	0.177094
CALLE	188	0.009930
CARRO	1262	0.365442
CARTON	238	0.026481
CASA	403	0.081099
CIUDAD	225	0.022178
CONTAMINAC...	1151	0.328699
CUIDADO	542	0.127110
ECOLOGIA	497	0.112214
ECOSISTEMA	811	0.216154
EDIFICIO	865	0.234028
ENFERMEDAD	254	0.031778
FABRICA	318	0.052963
FLORA	246	0.029129
GENTE	303	0.047997
GRANDE	229	0.023502
HUMO	506	0.115194
LATA	241	0.027474
LIMPIA	453	0.097650
NATURAL	1386	0.406488
OXIGENO	383	0.074479
PAPEL	553	0.130751
PERSONA	574	0.137703
PLANTA	2509	0.778219
PLASTICO	699	0.179080
RECICLAJE	544	0.127772
REUSAR	720	0.186031
RUIDO	229	0.023502
SMOG	345	0.061900
SUCIO	451	0.096988
TALA	158	0.000000
TRAFICO	253	0.031447
VERDE	1143	0.326051
VIDA	602	0.146971

APÉNDICE C

Matrices de comparación entre principiantes

Matrices de comparación generadas por el programa que compara entre las RSN de principiantes.

Tabla C. 1. Matriz de comparación entre ECOLOGÍA 1 v ECOLOGÍA

	ECO 1	ECO 3
AGUA	0.166	0.032
AIRE	0.084	0.233
ALUMINIO		0.004
AMBIENTE	0.371	0.397
ANIMAL	0.562	0.761
ARBOL	1.000	1.000
AYUDA		0.066
BASURA	0.636	0.704
BOSQUE	0.269	0.081
BOTE	0.020	
BOTELLA	0.020	
CALLE	0.026	
CARRO	0.381	0.420
CASA	0.135	0.048
CIUDAD		0.029
CLIMA		0.031
CONTAMINACION	0.273	0.585
CUIDADO	0.030	0.205
DESECHO	0.033	
ECOLOGIA	0.267	
ECOSISTEMA	0.191	0.185
EDIFICIO	0.305	0.215
EMPRESA		0.031
ENFERMEDAD	0.071	0.066
FABRICA	0.015	0.126
FAUNA	0.038	0.036
FLORA	0.069	
GENTE	0.038	0.115
GRANDE		0.025
HUMO	0.125	0.260
LATA	0.036	0.029
LIMPIA	0.072	
MEDIO		0.031
METROPOLI	0.030	
MONTANA		0.061
NATURAL	0.316	0.465
NECESIDAD		0.001
ORGANICO	0.031	
OXIGENO	0.087	0.196
PAPEL	0.110	0.086
PERSONA	0.171	0.106
PLANTA	0.794	0.636
PLASTICO	0.152	0.131
RECICLAJE	0.125	0.142
REUSAR	0.278	0.136
RIO	0.036	0.027
RUIDO		0.056
SEMILLA	0.020	
SMOG	0.063	0.099
SUCIO	0.112	0.068
TALA		0.007
TIENDAS	0.001	
TIERRA	0.041	
TOXICO	0.001	
TRAFICO	0.081	0.022
VERDE	0.493	0.346
VIDA	0.049	0.203

Tabla C. 2. Matriz de comparación entre ECOLOGÍA 2 y ECOLOGÍA 3

	ECO 2	ECO 3
AGUA	0.359	0.032
AIRE	0.181	0.233
ALUMINIO	0.139	0.004
AMBIENTE	0.213	0.397
ANIMAL	0.840	0.761
ARBOL	0.965	1.000
AYUDA		0.066
BASURA	0.518	0.704
BOSQUE	0.174	0.081
CARRO	0.375	0.420
CARTON	0.092	
CASA	0.084	0.048
CIUDAD		0.029
CLIMA		0.031
CONTAMINACIÓN	0.306	0.585
CUIDADO	0.055	0.205
ECOSISTEMA	0.149	0.185
EDIFICIO	0.164	0.215
EMPRESA	0.015	0.031
ENFERMEDAD		0.066
FABRICA	0.039	0.126
FAUNA		0.036
GENTE	0.039	0.115
GRANDE	0.020	0.025
HUMO	0.076	0.260
INORGANICO	0.031	
LATA	0.041	0.029
LIMPIA	0.140	
MEDIO		0.031
MONTAÑA		0.061
NATURAL	0.383	0.465
NECESIDAD		0.001
OXIGENO	0.027	0.196
PAPEL	0.168	0.086
PERSONA	0.154	0.106
PLANTA	1.000	0.636
PLÁSTICO	0.268	0.131
RECICLAJE	0.138	0.142
REUSAR	0.191	0.136
RIO		0.027
RUIDO	0.074	0.056
SMOG	0.037	0.099
SUCIO	0.117	0.068
TALA	0.001	0.007
TIERRA	0.264	
TRAFICO	0.015	0.022
VERDE	0.463	0.346
VIDA	0.130	0.203
VIDRIO	0.029	

Tabla C. 3. Matriz de comparación entre ECOLOGÍA 2 y ECOLOGÍA 4.

	ECO 2	ECO 4
AGUA	0.359	0.327
AIRE	0.181	0.235
ALUMINIO	0.139	0.057
AMBIENTE	0.213	0.195
ANIMAL	0.840	0.598
ARBOL	0.965	1.000
BASURA	0.518	0.379
BOSQUE	0.174	0.052
BOTELLA		0.024
CALLE		0.014
CARRO	0.375	0.292
CARTÓN	0.092	
CASA	0.084	0.043
CIUDAD		0.054
CONTAMINACIÓN	0.306	0.373
CUIDADO	0.055	0.059
ECOLOGÍA		0.030
ECOSISTEMA	0.149	0.224
EDIFICIO	0.164	0.241
EMPRESA	0.015	
ENFERMEDAD		0.022
FABRICA	0.039	0.027
FLORA		0.014
GENTE	0.039	0.004
GRANDE	0.020	0.022
HUMO	0.076	0.041
INORGANICO	0.031	
LATA	0.041	
LIMPIA	0.140	0.009
MEDIO		0.014
NATURAL	0.383	0.432
OXÍGENO	0.027	0.067
PAPEL	0.168	0.114
PERSONA	0.154	0.098
PLANTA	1.000	0.621
PLÁSTICO	0.268	0.121
POBLACION		0.057
RECICLAJE	0.138	0.090
REUSAR	0.191	0.130
RUIDO	0.074	0.017
SMOG	0.037	0.042
SUCIO	0.117	0.065
TALA	0.001	0.001
TIERRA	0.264	0.138
TRAFICO	0.015	
VERDE	0.463	0.205
VIDA	0.130	0.208
VIDRIO	0.029	

Tabla C. 4. Matriz de comparación entre ECOLOGÍA 3 y ECOLOGÍA 4.

	ECO 3	ECO 4
AGUA	0.032	0.327
AIRE	0.233	0.235
ALUMINIO	0.004	0.057
AMBIENTE	0.397	0.195
ANIMAL	0.781	0.598
ARBOL	1.000	1.000
AYUDA	0.066	
BASURA	0.704	0.379
BOSQUE	0.081	0.052
BOTELLA		0.024
CALLE		0.014
CARRO	0.420	0.292
CASA	0.046	0.043
CIUDAD	0.029	0.054
CLIMA	0.031	
CONTAMINACION	0.585	0.373
CUIDADO	0.205	0.059
ECOLOGIA		0.030
ECOSISTEMA	0.185	0.224
EDIFICIO	0.215	0.241
EMPRESA	0.031	
ENFERMEDAD	0.066	0.022
FABRICA	0.126	0.027
FAUNA	0.036	
FLORA		0.014
GENTE	0.115	0.004
GRANDE	0.025	0.022
HUMO	0.260	0.041
LATA	0.029	
LIMPIA		0.009
MEDIO	0.031	0.014
MONTANA	0.061	
NATURAL	0.465	0.432
NECESIDAD	0.001	
OXIGENO	0.196	0.067
PAPEL	0.086	0.114
PERSONA	0.106	0.098
PLANTA	0.636	0.621
PLASTICO	0.131	0.121
POBLACION		0.057
RECICLAJE	0.142	0.090
REUSAR	0.136	0.130
RIO	0.027	
RUIDO	0.056	0.017
SMOG	0.099	0.042
SUCIO	0.068	0.065
TALA	0.007	
TALA		0.001
TIERRA		0.138
TRAFICO	0.022	
VERDE	0.346	0.205
VIDA	0.203	0.208

APÉNDICE D

Matrices de comparación entre Red Total de Expertos y Principiantes

Matrices de comparación generadas por el programa que compara entre cada proyecto de principiantes y la Red Total de Expertos.

Tabla D. 1. Matriz de Comparación entre ECOLOGÍA 1 y la Red Total de Expertos.

	ECO 1	RED DE EXPERTOS		ECO 1	RED DE EXPERTOS
AGUA	0.166	0.028	PLANETA		0.250
AIRE	0.084	0.056	PLANTA	0.794	0.250
AMBIENTE	0.371		PLASTICO	0.152	
ANIMAL	0.562		POBLACION		1.000
APROVECHAR		0.306	PROCESAR		0.056
ARBOL	1.000	0.333	QUIMICOS		0.001
AREA VERDE		0.056	RECICLAJE	0.125	
ARMONIA		0.056	RECUPERAR		0.139
BASURA	0.638	0.894	RELACION		0.139
BIODIVERSIDAD		0.333	RESIDUO		0.028
BIOSFERA		0.194	REUSAR	0.278	
BOSQUE	0.269	0.333	REUTILIZAR		0.056
BOTE	0.020		RIO	0.038	
BOTELLA	0.020		RURAL		0.056
CALLE	0.026		SEMILLA	0.020	
CARRO	0.381	0.583	SMOG	0.063	0.417
CASA	0.135	0.444	SUCIO	0.112	
CIENCIA		0.056	SUELO		0.306
CONCIENCIA		0.056	SUSTENTABILIDAD		0.333
CONTAMINACION	0.273	0.056	TIENDAS	0.001	
CONTAMINANTES		0.056	TIERRA	0.041	
CUIDADO	0.030	0.028	TOXICO	0.001	
DESCOMPOSICION		0.028	TRAFICO	0.081	0.001
DESECHO	0.033	0.028	TRANSFORMACION		0.028
DESTRUCCION		0.028	UNIVERSO		0.028
DETERIORO		0.056	URBANO		0.083
ECOLOGIA	0.267		VEGETACION		0.500
ECOSISTEMA	0.191		VERDE	0.493	0.417
EDIFICIO	0.305	0.389	VIDA	0.049	0.222
EMPRESA		0.056	RECUPERAR		0.250
ENFERMEDAD	0.071		RELACION	0.794	0.250
EQUILIBRIO		0.333	RESIDUO	0.152	
ESTUDIO		0.111	REUSAR		1.000
FABRICA	0.015	0.001	REUTILIZAR		0.056
FAUNA	0.038	0.111	RIO		0.001
FLORA	0.069	0.167	RURAL	0.125	
GAS		0.056	SEMILLA		0.139
GENERADOR		0.001	SMOG		0.139
GENTE	0.038	0.056	SUCIO		0.028
HABITAT		0.139	SUELO	0.278	
HOMBRE		0.417	SUSTENTABILIDAD		0.056
HUMO	0.125		TIENDAS	0.038	
INDUSTRIA		0.028	TIERRA		0.056
INTERACCION		0.333	TOXICO	0.020	
LATA	0.038		TRAFICO	0.063	0.417
LIMPIA	0.072		TRANSFORMACION	0.112	
METROPOLI	0.030	0.278	UNIVERSO		0.306
NATURAL	0.316		URBANO		0.333
ORGANICO	0.031		VEGETACION	0.001	
ORGANISMO		0.028	VERDE	0.041	
OXIGENO	0.087	0.028	VIDA	0.001	
OZONO		0.028			
PAISAJE		0.111			
PAPEL	0.110				
PERSONA	0.171				

Tabla D. 2. Matriz de Comparación entre ECOLOGÍA 2 y la Red Total de Expertos.

	ECO 2	RED DE EXPERTOS		ECO 2	RED DE EXPERTOS
AGUA	0.359	0.028	LATA	0.041	
AIRE	0.181	0.056	LIMPIA	0.140	
ALUMINIO	0.139		METROPOLI		0.278
AMBIENTE	0.213		NATURAL	0.383	
ANIMAL	0.840		ORGANISMO		0.028
APROVECHAR		0.306	OXIGENO	0.027	0.028
ARBOL	0.965	0.333	OZONO		0.028
AREA VERDE		0.056	PAISAJE		0.111
ARMONIA		0.056	PAPEL	0.168	
BASURA	0.518	0.694	PERSONA	0.154	
BIODIVERSIDAD		0.333	PLANETA		0.250
BIOSFERA		0.194	PLANTA	1.000	0.250
BOSQUE	0.174	0.333	PLÁSTICO	0.268	
CARRO	0.375	0.583	POBLACION		1.000
CARTÓN	0.092		PROCESAR		0.056
CASA	0.084	0.444	QUIMICOS		0.001
CIENCIA		0.056	RECICLAJE	0.138	
CONCIENCIA		0.056	RECUPERAR		0.139
CONTAMINACIÓN	0.306	0.056	RELACION		0.139
CONTAMINANTES		0.056	RESIDUO		0.028
CUIDADO	0.055	0.028	REUSAR	0.191	
DESCOMPOSICIÓN		0.028	REUTILIZAR		0.056
DESECHO		0.028	RUIDO	0.074	
DESTRUCCIÓN		0.028	RURAL		0.056
DETERIORO		0.056	SMOG	0.037	0.417
ECOSISTEMA	0.149		SUCIO	0.117	
EDIFICIO	0.164	0.389	SUELO		0.306
EMPRESA	0.015	0.056	SUSTENTABILIDAD		0.333
EQUILIBRIO		0.333	TALA	0.001	
ESTUDIO		0.111	TIERRA	0.264	
FABRICA	0.039	0.001	TRAFICO	0.015	0.001
FAUNA		0.111	TRANSFORMACIÓN		0.028
FLORA		0.167	UNIVERSO		0.028
GAS		0.056	URBANO		0.083
GENERADOR		0.001	VEGETACION		0.500
GENTE	0.039	0.056	VERDE	0.463	0.417
GRANDE	0.020		VIDA	0.130	0.222
HABITAT		0.139	VIDRIO	0.029	
HOMBRE		0.417			
HUMO	0.076				
INDUSTRIA		0.028			
INORGANICO	0.031				
INTERACCIÓN		0.333			

Tabla D. 3. Matriz de Comparación entre ECOLOGÍA 3 y la Red Total de Expertos.

	ECO 3	RED DE EXPERTOS		ECO 3	RED DE EXPERTOS
AGUA	0.032	0.028	INTERACCIÓN		0.333
AIRE	0.233	0.056	LATA	0.029	
ALUMINIO	0.004		MEDIO	0.031	
AMBIENTE	0.397		METROPOLI		0.278
ANIMAL	0.761		MONTAÑA	0.061	
APROVECHAR		0.306	NATURAL	0.465	
ARBOL	1.000	0.333	NECESIDAD	0.001	
AREA VERDE		0.056	ORGANISMO		0.028
ARMONIA		0.056	OXIGENO	0.196	0.028
AYUDA	0.066		OZONO		0.028
BASURA	0.704	0.694	PAISAJE		0.111
BIODIVERSIDAD		0.333	PAPEL	0.086	
BIOSFERA		0.194	PERSONA	0.106	
BOSQUE	0.081	0.333	PLANETA		0.250
CARRO	0.420	0.583	PLANTA	0.636	0.250
CASA	0.048	0.444	PLÁSTICO	0.131	
CIENCIA		0.056	POBLACION		1.000
CIUDAD	0.029		PROCESAR		0.056
CLIMA	0.031		QUIMICOS		0.001
CONCIENCIA		0.056	RECICLAJE	0.142	
CONTAMINACIÓN	0.585	0.056	RECUPERAR		0.139
CONTAMINANTES		0.056	RELACION		0.139
CUIDADO	0.205	0.028	RESIDUO		0.028
DESCOMPOSICIÓN		0.028	REUSAR	0.136	
DESECHO		0.028	REUTILIZAR		0.056
DESTRUCCIÓN		0.028	RIO	0.027	
DETERIORO		0.056	RUIDO	0.056	
ECOSISTEMA	0.185		RURAL		0.056
EDIFICIO	0.215	0.389	SMOG	0.099	0.417
EMPRESA	0.031	0.056	SUCIO	0.068	
ENFERMEDAD	0.066		SUELO		0.306
EQUILIBRIO		0.333	SUSTENTABILIDAD		0.333
ESTUDIO		0.111	TALA	0.007	
FABRICA	0.126	0.001	TRAFICO	0.022	0.001
FAUNA	0.036	0.111	TRANSFORMACIÓN		0.028
FLORA		0.167	UNIVERSO		0.028
GAS		0.056	URBANO		0.083
GENERADOR		0.001	VEGETACION		0.500
GENTE	0.115	0.056	VERDE	0.346	0.417
GRANDE	0.025		VIDA	0.203	0.222
HABITAT		0.139			
HOMBRE		0.417			
HUMO	0.260				
INDUSTRIA		0.028			

Tabla D. 4. Matriz de Comparación entre ECOLOGÍA 4 y la Red Total de Expertos.

	ECO 4	RED DE EXPERTOS		ECO 4	RED DE EXPERTOS
AGUA	0.327	0.028	INDUSTRIA		0.028
AIRE	0.235	0.056	INTERACCIÓN		0.333
ALUMINIO	0.057		LIMPIA	0.009	
AMBIENTE	0.195		MEDIO	0.014	
ANIMAL	0.598		METROPOLI		0.278
APROVECHAR		0.306	NATURAL	0.432	
ARBOL	1.000	0.333	ORGANISMO		0.028
AREA VERDE		0.056	OXIGENO	0.067	0.028
ARMONIA		0.056	OZONO		0.028
BASURA	0.379	0.694	PAISAJE		0.111
BIODIVERSIDAD		0.333	PAPEL	0.114	
BIOSFERA		0.194	PERSONA	0.098	
BOSQUE	0.052	0.333	PLANETA		0.250
BOTELLA	0.024		PLANTA	0.621	0.250
CALLE	0.014		PLÁSTICO	0.121	
CARRO	0.292	0.583	POBLACION	0.057	1.000
CASA	0.043	0.444	PROCESAR		0.056
CIENCIA		0.056	QUIMICOS		0.001
CIUDAD	0.054		RECICLAJE	0.090	
CONCIENCIA		0.056	RECUPERAR		0.139
CONTAMINACIÓN	0.373	0.056	RELACION		0.139
CONTAMINANTES		0.056	RESIDUO		0.028
CUIDADO	0.059	0.028	REUSAR	0.130	
DESCOMPOSICIÓN		0.028	REUTILIZAR		0.056
DESECHO		0.028	RUIDO	0.017	
DESTRUCCION		0.028	RURAL		0.056
DETERIORO		0.056	SMOG	0.042	0.417
ECOLOGÍA	0.030		SUCIO	0.065	
ECOSISTEMA	0.224		SUELO		0.306
EDIFICIO	0.241	0.389	SUSTENTABILIDAD		0.333
EMPRESA		0.056	TALA	0.001	
ENFERMEDAD	0.022		TIERRA	0.138	
EQUILIBRIO		0.333	TRAFICO		0.001
ESTUDIO		0.111	TRANSFORMACIÓN		0.028
FABRICA	0.027	0.001	UNIVERSO		0.028
FAUNA		0.111	URBANO		0.083
FLORA	0.014	0.167	VEGETACION		0.500
GAS		0.056	VERDE	0.205	0.417
GENERADOR		0.001	VIDA	0.208	0.222
GENTE	0.004	0.056			
GRANDE	0.022				
HABITAT		0.139			
HOMBRE		0.417			
HUMO	0.041				