

## Visualización de Redes Semánticas Naturales a través de Acomodo de Grafos Dirigido por Fuerzas

### *Natural Semantic Network Visualization by Means of a Graph Spring-Force Layout*

Torres Guerrero, Francisco & Garza, Sara E.

**Resumen.** En la presente investigación se realiza un estudio exploratorio con el objetivo de realizar un análisis para la creación de una visualización de una red semántica natural de la temática de ecología. Se realizaron tres diferentes visualizaciones utilizando las aplicaciones Knot (para la visualización de Pathfinder), Wordle (para la visualización de *Tag Cloud*) y *Smart Semantic Cloud* (para la visualización de *Semantic Cloud*). Este último método es el que proponemos.

**Abstract.** We perform an exploratory study for natural semantic network visualization. Three visualizations are generated, each with a different tool and method: Knot for the Pathfinder method, Wordle for the tag cloud method, and Smart Semantic Cloud for the Semantic Cloud method. This last method (as well as the tool for generating visualizations) is our proposal.

**Palabras clave.** Visualización, Redes Semánticas Naturales, Nube de Etiquetas, Pathfinder

**Keywords.** Visualization, Natural Semantic Networks, Tag Cloud, Pathfinder

### Introducción

Todo conocimiento se encuentra expresado por información y esta a su vez por datos que pueden ser expresados de diferentes maneras, por ejemplo a través de texto y gráficos. Podemos definir la representación del conocimiento como la visualización en la que se estructuran los datos que al unirse construyen información, y al formar un sistema se convierten en una representación del conocimiento (Herman 2006; Markman, 1999). Sin importar el formato de la visualización de una representación del conocimiento, esta visualización es procesada e interpretada por el cerebro; sin embargo, una representación bien graficada permite un mejor procesamiento de información (Ware, 2013). Las Redes Semánticas Naturales (RSN) constituyen una técnica de representación del conocimiento de una temática que permite estructurar datos obtenidos a través de las respuestas de los participantes de un grupo. Las redes semánticas naturales como herramienta de representación de conocimiento han sido ampliamente utilizadas para el estudio del significado por diferentes investigadores y áreas (Figueroa, González & Solís, 1981; López, 2002). Es de interés para las RSN's construir representaciones del conocimiento de un grupo que comparta un mismo perfil y de un mismo nivel de dominio del tema, ya que de esta manera se puede realizar un análisis de contenido más preciso; por ejemplo, la construcción de una red semántica realizada por nutriólogos sobre la obesidad nos brindará un panorama de conocimiento con un enfoque profesional (Figueroa, González & Solís, 1981; López, 2002). Aunque la técnica de RSN cuenta con una visualización del tipo matriz donde expresa sus valores importantes con formato numérico y texto no cuenta con herramientas graficas que permitan discriminar o deducir de manera óptima la información contenida. Existe la necesidad de seguir realizando un trabajo de investigación que permita la innovación en la visualización del mismo (De la Garza, Sánchez & López, 2010).

## Objetivos

En el presente estudio se realiza una investigación sobre la estructura de las redes semánticas naturales donde se evalúan dos métodos de visualización: Pathfinder y *Tag Cloud*. También se propone un tercer método que permita la evaluación de una red semántica, el cual denominaremos como *Semantic Cloud*. Actualmente existen investigaciones donde se evalúa utilizando la técnica de Pathfinder a través de la aplicación Knot como instrumento para visualizar RSN; por tal motivo, el objetivo es comparar y estudiar dos visualizadores adicionales con un mayor alcance en opciones gráficas. El presente estudio es una investigación con un enfoque exploratorio.

Se plantean las siguientes preguntas de investigación:

¿Se puede utilizar el método *Tag Cloud* para realizar una visualización efectiva de las RSN?

¿Se puede utilizar el método *Semantic Cloud* para realizar una visualización efectiva de las RSN?

## Marco Teórico

### *Redes Semánticas Naturales*

La técnica de RSN se construye a partir de las respuestas de un grupo de participantes con un perfil en común acerca de una temática definida. Debido a que las respuestas son ponderadas por los mismos participantes se pueden realizar los cálculos que permiten identificar los conceptos más relevantes y descartar los menos importantes. De una RSN son de relevancia los siguientes indicadores: El valor J que es la cantidad de conceptos diferentes que constituyen la red; el valor G indica la densidad o dispersión de la red, esto con respecto al concepto central; y el grupo SAM. Cada grupo SAM se compone de los diez conceptos que son más relevantes al estímulo en el cual se puede observar que se desprende el valor F que corresponde al número incidencias del concepto en la red y el M que es un indicador de la ponderación dada por los participantes (Figuroa, González & Solís, 1981; López, 2002).

### *Pathfinder*

La técnica Pathfinder toma las estimaciones de las proximidades entre pares de elementos como entrada y construye una red semántica. La red pathfinder utiliza los conceptos como nodos y un conjunto de nodos para las estimaciones de proximidad simétrica o asimétrica que conectan pares de nodos. Los enlaces están definidos a través de patrones de proximidad de los conceptos y parámetros de los algoritmos de Pathfinder (Schvaneveldt, 1990). Actualmente se puede construir una red pathfinder utilizando la herramienta Knot—por sus siglas en inglés *Knowledge Network Organizing Tool* (Interlink, 2013).

### *Tagcloud Wordle*

El método *Tag Cloud* o *Nube de Etiquetas* es una técnica que permite representar conceptos de una temática mediante la creación de una visualización. Las nubes

pretenden captar la atención del usuario mediante la manipulación propiedades visuales como lo es el acomodo, tamaño, color, posición y tipo de letra (Bateman, Gutwin & Nacenta 2008).

### *Visualización Semantic Cloud*

La visualización que a continuación se describe pretende contribuir al análisis de redes semánticas naturales, y es nuestra propuesta. Esta visualización considera a una red semántica natural como un grafo bipartito *estímulos-definidores* y asume que la proyección *definidor-definidor* del grafo es suficientemente intuitiva para ilustrar la red. La técnica que se sigue está inspirada en *nubes semánticas de etiquetas*, las cuales hacen un acomodo de elementos que no es alfabético o estético, sino más bien basado en la cercanía de una etiqueta con respecto a otra. En nuestro caso, cada definidor es una etiqueta de la nube y su tamaño de letra está dado por un índice de centralidad o importancia; este índice se calcula con base en el valor F. Por otra parte, la cercanía entre definidores está determinada por su co-ocurrencia en uno o más grupos SAM. El acomodo se realiza con un algoritmo para dibujar grafos que está dirigido por fuerzas (Fruchterman & Reinglod, 1991). La nube se complementa con colores y puntos que ilustran los grupos SAM de la red.

### *Representación de la red*

La red semántica natural puede verse como un *grafo bipartito*. En un grafo bipartito, por definición, existen dos tipos (disjuntos) de vértices y cada arista conecta siempre pares de vértices que son de tipos distintos. En una RSN, los estímulos corresponderían a un tipo de vértice y los definidores a otro. Una arista  $\{d, e\}$  implicaría la existencia de un definidor  $d$  en el grupo SAM del estímulo  $e$ .

Del grafo bipartito, se utiliza la proyección *definidor-definidor* para generar la visualización. Esto es por dos motivos. En primer lugar, es común que el análisis de grafos bipartitos se haga por proyecciones, pues cada proyección es un grafo más sencillo de manejar y estudiar. En segundo lugar, consideramos que los definidores son la parte que realmente enriquece a la red; por tanto, sería de mayor provecho visualizar una RSN a través de sus definidores. Al realizar esta proyección, lógicamente todos los vértices son definidores; una arista indica, en este caso, que un par de estos aparece en el mismo grupo SAM. Nótese que el grafo es no-ponderado; por lo tanto, cada arista tiene peso de 1.

### *Etiquetas (texto)*

Cada definidor del grafo proyectado se puede ver como una etiqueta (fragmento de texto) perteneciente a una *nube de etiquetas*. Como sabemos, en una nube cada etiqueta tiene un tamaño que es proporcional a su importancia para el contexto con el que se esté trabajando. Para obtener el tamaño de cada definidor, utilizamos un sencillo índice de centralidad. El índice está dado por el valor F del definidor, puesto que tal valor sirve para detectar términos nucleares en la red. Para cada definidor, entonces, se toma su F y se relaciona (“mapea”) a un tamaño de letra específico; para evitar tamaños muy grandes o muy pequeños que pudiesen afectar la visualización, establecemos un rango de [10, 30]. Para generar tamaños de letra que caigan en este rango, utilizamos la

Ecuación 1:

$$\left| \frac{f_i}{f_{\max(\beta \alpha + 1)}} \right| + \alpha$$

donde  $f_i$  es el valor F del definidor,  $f_{\max}$  es el valor F más grande de la red y, en nuestro caso,  $\alpha=10$  y  $\beta=30$ . Si todas las etiquetas tienen la misma frecuencia, el tamaño se fija en 10 para evitar que la nube se vea sobrecargada.

Para generar la nube semántica, se echa mano de un algoritmo de acomodo de grafos dirigido por fuerzas. Este tipo de algoritmos se basa en física y genera fuerzas atractivas para vértices que comparten aristas (o aristas fuertes, en el caso de grafos ponderados) y fuerzas repulsivas en el caso contrario.

### Colores

Cada grupo SAM tiene su propio color dentro de la visualización; estos se escogen automáticamente de una gama predefinida. Si un definidor se encuentra en más de un grupos SAM, su color pasa a ser el negro (#000000 en hexadecimal).

### Nubes de puntos

Para apreciar el valor M de los definidores, en la visualización se incluyen “nubes de puntos” dentro del área de cada definidor. Entre mayor es el valor M, más densa es esta nube, i.e. más puntos se colocan para el definidor. Estos puntos son de un color parecido (o el mismo color, pero diferente tono) al del grupo SAM donde se encuentra el definidor. En caso de que el definidor se halle en más de un grupo, el área interna se subdivide para poder colocar los puntos pertenecientes a cada grupo.

### Método

En este estudio, a través de una técnica de redes semánticas naturales, se obtuvieron de los participantes definiciones conceptuales de siete conceptos centrales al esquema de medio ambiente (Eco-esquema) a través de la aplicación conocida como SemNet (De la Garza, Sánchez & López, 2010). Una vez obtenida la red semántica, se procedió a realizar las visualizaciones utilizando las aplicaciones Knot para la visualización de Pathfinder, Wordle para la visualización de *Tag Cloud* y *Smart Semantic Cloud* (aplicación escrita en Python) para la visualización de *Semantic Cloud*.

### Resultados

En la presente investigación, se procedió a trabajar con una red de la temática *ecología*, la cual fue construida mediante la colaboración de 40 participantes, los cuales son estudiantes de universitarios del área de ingeniería. Esta red semántica se muestra en la figura 1.

CIUDAD		CONTAMINACION		ECOLOGIA		MEDIO AMBIENTE		NATURALEZA			
F	Definidor	M	F	Definidor	M	F	Definidor	M	F	Definidor	M
2	CARROS	325	1	HUMO	263	3	RECICLAR	151	4	NATURALEZA	186
1	EDIFICIOS	205	2	BASURA	241	4	AMBIENTE	150	3	AGUA	127
2	CIUDAD	195	2	CARROS	177	4	NATURALEZA	128	3	ANIMALES	112
3	CONTAMINACION	168	3	CONTAMINACION	152	3	ANIMALES	128	2	ECOSISTEMAS	106
1	CALLES	160	2	FABRICAS	113	2	PLANTAS	117	3	CONTAMINACION	105
1	CASAS	148	3	AGUA	112	4	ARBOLES	113	4	ARBOLES	103
1	PERSONAS	139	2	CIUDAD	78	2	ECOSISTEMAS	81	3	AIRE	102
1	GENTE	138	3	AIRE	71	2	VERDE	77	1	TIERRA	81
2	FABRICAS	121	2	RUIDO	61	4	ECOLOGIA	77	1	CALENTAMIENTO	73
2	RUIDO	62	4	AMBIENTE	58	1	RECICLAJE	62	4	ECOLOGIA	57
J = 93		J = 83		J = 77		J = 74		J = 71			
G = 26.3		G = 20.5		G = 8.9		G = 12.9		G = 19.7			

RECICLAJE		REFORESTACION									
F	Definidor	M	F	Definidor	M	F	Definidor	M	F	Definidor	M
1	PAPEL	220	4	ARBOLES	333						
2	BASURA	171	2	BOSQUES	167						
1	PLASTICO	152	1	TALA	109						
1	ALUMINIO	145	3	RECICLAR	89						
3	RECICLAR	133	1	VIDA	84						
1	LATAS	95	4	NATURALEZA	80						
1	CARTON	70	2	FLORA	74						
1	VIDRIO	69	2	VERDE	73						
4	AMBIENTE	65	1	PLANTAR	73						
4	ECOLOGIA	48	4	ECOLOGIA	67						
J = 77		J = 73									
G = 17.2		G = 26.6									

Total Conceptos	7	Def. x Concep.	284	Conceptos Diferentes	38	H-S 1	H-S 2	Compara	Predictor
Total Alumnos	40.0	Def. x Alum.	49.8	Conceptos Comunes	19(51)				
Total Definidores	1,990	Def. x Alum.x Concep	7.1	Conceptos Libres	19				

Figura 1. Pantalla semnet con los resultados del estudio de redes.

El orden de conectividad para los grupos SAMs de redes semánticas fue el siguiente:

Animales (4)    Contaminación (3),    Aire (3),    Carros (2),    Naturaleza (2),  
 Ruido (2),  
 Árboles (4)    Plantas (3),    Tierra (3),    Ecosistema (2),    Vida (2)  
 Verde (3),  
 Agua (4)    Basura (3),

Todos los demás definidores solo estuvieron presentes en su propio grupo SAM.

Cuando se procesa el análisis Pathfinder para el caso de la red conceptual de los usuarios se obtiene una organización de proximidad semántica. En la figura 2 se muestra la visualización Pathfinder correspondiente a RSN de ecología.

Para poder realizar la visualización utilizando *Tag Cloud*, se extrajeron todas las palabras de los diferentes grupos SAM. Se utilizaron las propiedades visuales de tamaño de texto para indicar los conceptos con mayor incidencia. La aplicación utiliza un color diferente para cada una de los conceptos. Los conceptos aparecen agrupados en el centro eliminando grandes espacios en blanco.

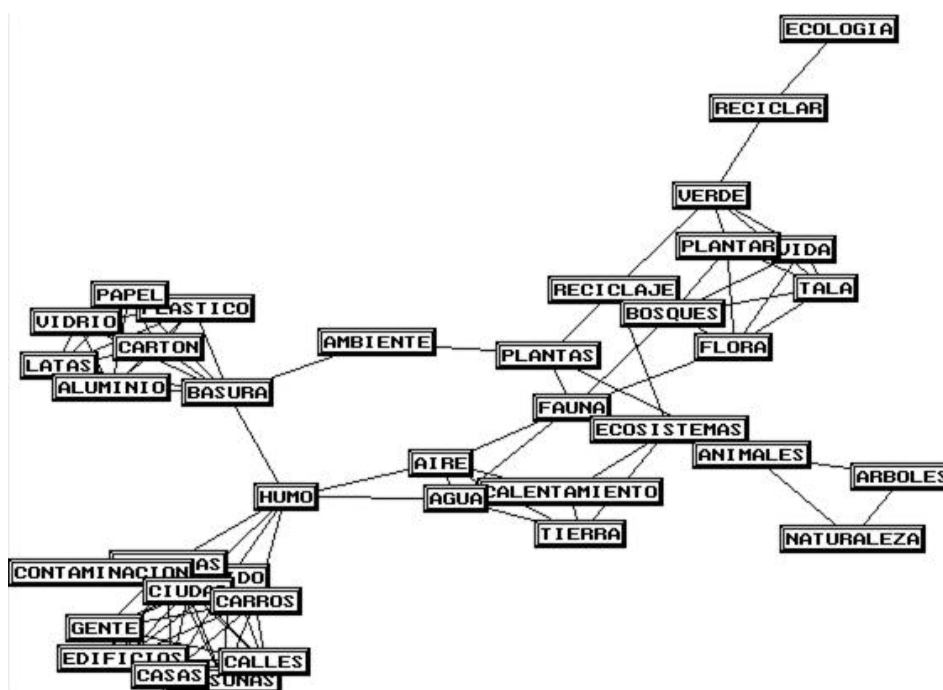
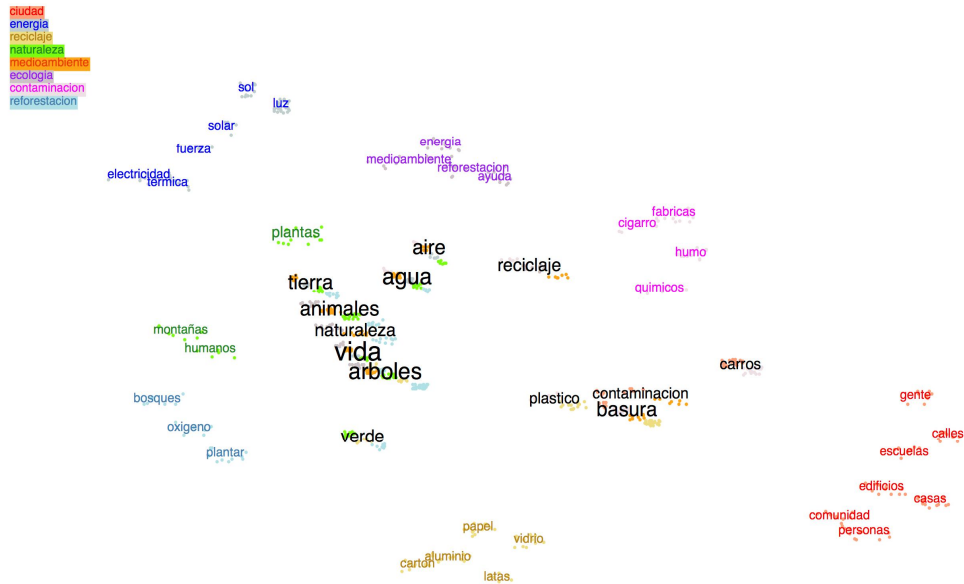


Figura 2. Visualización de la RSN de ecología a través técnica Pathfinder utilizando la aplicación KNOT



Figura 3. Visualización de la RSN de ecología a través técnica Tag Cloud utilizando la aplicación Wordle

La visualización del grafo dirigido por fuerzas toma en cuenta tres aspectos principales de la RSN que corresponden a: el número de incidencias del concepto de los diferentes grupos SAM, la distinción del concepto y el grupo al que pertenece, y el valor M el cual se expresa por la posición visual espacial el cual a mayor valor aparece más cerca al punto central del grafo. En la figura 4 se visualiza la red semántica de ecología



**Figura 4.** Visualización de la RSN de ecología a través técnica *Semantic Cloud* utilizando la aplicación *Smart Semantic Cloud*.

## Conclusiones

La visualización de una red semántica puede tener diferentes alcances a partir de la técnica empleada para representarla. Se pueda observar que en las visualizaciones empleadas tienen diferente cantidad de indicadores pertenecientes a la RSN los cuales se distribuyen de la siguiente manera: *Semantic Cloud* muestra seis indicadores que son definidores, conexiones entre conceptos, distancia semántica, grupo SAM, valor F y M, la visualización realizada mediante *Tag Cloud* muestra dos indicadores que son los definidores y valor F y por último la visualización construida por medio de *Pathfinder* muestra dos indicadores que son los definidores y proximidad. Si evaluamos las visualizaciones por la cantidad de información que muestra relevante a la RSN encontramos que *Semantic Cloud* cuenta con la mayor cantidad de indicadores. Es importante considerar que el objetivo principal de una visualización es que el usuario pueda identificar de manera apropiada la mayor información posible. Las propiedades visuales que son manipuladas por cada visualización pudieran ser elementos a favor o en contra para captar la atención del usuario. El exceso de información genera ruido lo que causa confusión y la falta información no proporciona las herramientas para la toma de decisiones. Como trabajo futuro se propone realizar un estudio de interacción humana

computadora para evaluar que visualización se ajusta mas al análisis de redes semánticas el cual se propone realizar herramientas como el *eye tracker*.

### Referencias

- De la Garza, G. A., Sánchez, M.M.P. & López, R. E. O (2010). SemNet.  
<http://www.cognilab.org/RECURSOS.html>.
- Figuroa, J. G., González, G. E., & Solís, V. M. (1981). Una aproximación al problema del significado  
Revista Latinoamericana de Psicología, 13(3), 447-458.
- Fruchterman T, Reingold E (1991) Graph drawing by force-directed placement. Software: Practice and  
experience 21(11):1129–1164.
- Jonathan Feinberg (2013) Wordle.  
<http://www.wordle.net/> . Consultado el: 10 diciembre 2014
- Hermann Helbig (2006) Knowledge Representation and the Semantic of Natural Language, Springer
- Interlink (2012) KNOT  
<http://interlinkinc.net/KNOT.html> . Consultado el: 10 diciembre 2014
- López, R.E.O. & Theios J. (1992). Semantic Analyzer of Schemata Organization (SASO). Behavior  
Research Methods, Instruments & Computers, 24(2), 277-285.
- López, R.E.O., Morales, M.G.E & Hedlefs, A.M.I (2011). El enfoque cognitivo de la formación de  
nuestros significados. En prensa México D.F.: Trillas.
- Markman, A.B. (1999) Knowledge Representation, Erlbaum
- Schaveneldt, R.W. (1990). Pathfinder Associative Networks: Studies in Knowledge Organization. New  
Jersey: Ablex.
- Valdez, J. (1998). Las redes semánticas naturales, usos y aplicaciones en psicología social. , Universidad  
Autónoma del Estado de México, 2, 65-80.
- Ware, C.(2013) Information visualization : perception for design. Amsterdam : Elsevier

---

Francisco Torres Guerrero & Sara E. Garza  
Universidad Autónoma de Nuevo León (UANL) FIME  
San Nicolás de los Garza, Nuevo León Mexico