

# ESTRUCTURACIÓN DE LA INFORMACIÓN BIBLIOGRÁFICA EN LA BASE CAB DE AUTORES VINCULADOS A INSTITUCIONES COLOMBIANAS, 1997 A 2002

## Structuring Bibliographic Information of Authors Linked with Colombian Institutions Indexed in CAB Database, 1997-2002

JORGE CHARUM, DIANA USGAME  
Observatorio Colombiano de Ciencia y Tecnología.

Presentado noviembre 30 de 2004, aceptado febrero 15 de 2005, correcciones abril 12 de 2005.

### RESUMEN

La utilización de la bibliometría para construir representaciones de las dinámicas de la producción científica y tecnológica se apoya en el tratamiento de las diferentes informaciones contenidas en las publicaciones científicas, sobre todo en los artículos, las patentes y los informes de investigación. La información recogida en la base bibliográfica CAB de los documentos integrados de 1997 a 2002, con autores vinculados a instituciones u organizaciones colombianas, ha sido estructurada para hacer emerger los temas de investigación presentes y para establecer los contextos en que se ubican los temas identificados, así como sus evoluciones. Se señalan diversas utilizaciones de la información estructurada.

**Palabras clave:** bibliometría, palabras asociadas, estructuración de la información, base bibliográfica, Colombia.

### ABSTRACT

The utilization of bibliometry in order to build representations of the scientific and technological production dynamics is supported on the treatment of the different information contained in scientific publications, above all in the articles, patents and investigation reports. The information gathered in the bibliographic base CAB of the documents integrated from 1997 to 2002, having authors linked with Colombian institutions or organizations, has been structured to make the present investigation themes emerge and in order to establish the contexts in which the themes identified as well as its evolutions are placed. Diverse utilizations of the structured information are noted.

**Key words:** bibliometry, co-words, structuring information, bibliographic base, Colombia.

### INTRODUCCIÓN

Una de las perspectivas generales de análisis de las dinámicas de la producción científica y tecnológica toma como referente a las diversas informaciones sobre las publi-

caciones científicas, sobre todo en los artículos, las patentes y los informes de investigación. El problema de la construcción de representaciones de esas dinámicas ha sido enfrentado desde diversas perspectivas teóricas, de las que se derivan procedimientos o algoritmos, por lo regular incorporados en programas de computación o *software*. La representación más inmediata corresponde a la elaboración de inventarios de documentos producidos, clasificados por ejemplo, por su tipo, por el área de la ciencia y por autores; bajo la hipótesis de que las cantidades permiten calificar el esfuerzo nacional en investigación.

A mediados de los años 60 del siglo pasado, Derek John de Solla Price (De Solla Price, 1965) propuso que se consideraran las redes de artículos científicos formadas por las citaciones que en cada uno de ellos se hacen a otros artículos, y se usara la estadística como herramienta para encontrar los diversos comportamientos regulares presentes en esas redes. Esta opción puede considerarse como el acta de nacimiento de la bibliometría y dio lugar a una gran cantidad de estudios relacionados con la producción de países, de áreas de la ciencia, entre otros, generando indicadores unidimensionales de actividad (Polanco, 1995). Las posibilidades abiertas por las nuevas tecnologías informáticas, que permitieron crear modalidades de acumulación de grandes cantidades de datos, y disminuir el tiempo y los costos de las operaciones que ponen en relación los datos sobre las citaciones, hicieron posible que, de manera centralizada, se dispusiera de todas las citaciones de un gran número de revistas a través del tiempo, y se establecieran jerarquías entre esas revistas con base en el cálculo de su factor de impacto para cada año<sup>1</sup>. El centro del análisis deja de ser el documento y deviene la revista. Se podría entonces conocer la cantidad y la calidad de la producción nacional identificando las publicaciones que aparecen en las revistas pertenecientes a un núcleo privilegiado de revistas con el mayor factor de impacto en sus especialidades; tal núcleo sería el vehículo de la llamada ciencia de la corriente principal.

Una nueva opción de análisis fue concebida considerando las cocitaciones (Small y Sweeney, 1985) presentes en los artículos de las revistas. Una cocitación es la presencia en un artículo de citaciones a otros dos artículos: si muchos autores coinciden en citar también de manera conjunta a ese par de artículos, se puede conjeturar que esta asociación es significativa y está en el centro de los intereses de los investigadores. De esta manera, la identificación del conjunto de artículos cocitados, y la introducción de medidas para hacer su jerarquización según el número de cocitaciones, permite la identificación del conjunto núcleo de artículos para cada especialidad de la ciencia, su *core set*, que son referencia casi obligada para los investigadores en esa especialidad y los frentes de investigación, formados por las publicaciones que cocitan

---

<sup>1</sup>El Institute for Scientific Information, ISI, desde comienzos de los años 70 creó grandes bases de datos siguiendo las propuestas de Price, analiza las citas que se hacen entre las revistas para elaborar una visión del desarrollo de la investigación en campos determinados. El factor de impacto establece la frecuencia promedio de citaciones a un artículo en una revista, y se calcula dividiendo el número de todas las citaciones a los artículos de una revista durante los dos años anteriores a un año determinado por el número de artículos publicados durante esos mismos dos años. Los datos sobre el factor de impacto son publicados anualmente.

los artículos de ese núcleo. La posibilidad de construir representaciones sobre las dinámicas de investigación nacional, pasa por la identificación de las publicaciones que pertenecen al núcleo y a los frentes de investigación con autores vinculados a instituciones u organizaciones nacionales.

Ya sea el análisis de las citas, que privilegia los artículos para derivar cualidades de las revistas, o el análisis de las cocitaciones, que lo hace para identificar conjuntos de artículos con cualidades derivadas, no tienen en cuenta sus contenidos. Para hacerlo se debe enfrentar entonces el problema de analizar datos textuales. Dos grandes orientaciones generales han estado en la base de la concepción de métodos para hacerlo. La primera, con una opción cuantitativa, proviene del análisis factorial o de las estadísticas; la segunda, con una opción cualitativa, integra elementos de la sociología de la ciencia, de la lingüística, así como también de la estadística (Leydesdorff, 1987). Las unidades básicas de análisis en esta opción son las palabras clave identificadas, representativas de conceptos o temáticas presentes dentro de un *corpus* documental. Así, un primer nivel del análisis textual se ocupa de la identificación de conceptos presentes en grandes conjuntos de documentos que permiten, en un segundo nivel, la construcción de agregaciones o *clusters* (Lejeune, 2004). En la siguiente tabla se ha hecho una síntesis de las opciones analíticas anteriores.

Unidades de análisis	Hipótesis de base	Forma de representación de las dinámicas de la investigación	Programas de computación
Artículos	La producción de artículos es representativa de la actividad	Inventarios y conteos	Programas genéricos para realizar operaciones aritméticas
Citaciones	Los artículos generan redes de citas	Núcleos significativos de autores, de revistas	Programas de base estadística
Cocitaciones	Las cocitaciones identifican asociaciones entre autores de artículos diferentes	Frentes de investigación	Programas de base estadística
Palabras clave o descriptores	Las palabras clave asociadas a un documento permiten una representación de su contenido	Redes de palabras: a) Métodos cuantitativos: métodos estadísticos multivariados b) Métodos cualitativos de análisis de contenido	a) SAS, SPSS, R, Téralogie, Spad.t, Neurodoc, T Lab b) Candide, Leximappe, T Lab, Réseau lu, Sdoc

Tabla 1. Niveles de representación de las dinámicas de la investigación basados en documentos y herramientas utilizadas para construirlos.

La utilización de los métodos citados anteriormente sirven para representar en diferentes niveles las dinámicas de la investigación, para identificar los investigadores: individuales y colectivos que contribuyen a ellas, y las instituciones u organizaciones a las que

pertenecen. La aplicación de los métodos se hace sobre los datos primarios. Toda aplicación da lugar a resultados y entonces, reto para los especialistas, es necesaria una interpretación cuidadosa que ponga en juego sus conocimientos específicos y la comprensión básica de los métodos y procedimientos utilizados. Es preciso entonces, de una parte conocer las nociones, conceptos y teorías que han dado lugar a esos métodos, y de otra conocer las formas que toman los datos primarios, sus atributos y las relaciones que pueden tener entre sí (Charum, 2004). Diferentes modalidades de análisis de la información bibliográfica han sido desarrolladas en las últimas décadas. El interés ha cambiado desde que el problema de la recuperación de la información estaba en el centro de las preocupaciones de los investigadores. En efecto, la construcción de bases bibliográficas que almacenan la información ha estado acompañada del diseño de interfaces que han banalizado las consultas y la extracción de los datos acumulados. El interés por el análisis de la información para tener representaciones sobre las dinámicas de la investigación pasó a ser determinante. Para ello la aplicación de nuevas teorías ha dado lugar a la concepción de nuevos instrumentos que, más allá de la recuperación de los datos primarios, pretenden extraer el conocimiento presente en las grandes bases bibliográficas. Se ha pasado de los datos a la identificación de relaciones entre los datos y al análisis del contenido, o dicho en otra forma, se ha pasado de los datos a la elaboración de la información, a la identificación del conocimiento presente en esas bases de datos, consideradas ahora como minas que deben ser explotadas.

La estructuración de la información se concibe como la construcción de información elaborada a partir de los datos primarios, organizada para responder a diferentes necesidades: de usuarios específicos, de comunidades de práctica, de elaboración de clasificaciones, y de identificación de regularidades y tendencias presentes en las bases de datos. Estas estructuraciones pueden hacerse: a) por la generación de información con base en las relaciones entre los datos; b) por la aplicación de teorías (sociológicas, estadísticas, lingüísticas, ...) que, además de orientar la generación de instrumentos, producen resultados que pueden ser interpretados desde esas mismas teorías; c) por la estructuración *ad hoc* para responder a las necesidades o problemas de usuarios o de comunidades de práctica explícitamente identificados. Cada uno de esos niveles de estructuración da lugar a usos específicos.

El análisis de las palabras asociadas (*co-word occurrence analysis*) ha sido ampliamente utilizado de diversas maneras en la investigación sobre los dominios del contenido que se encuentran en las bases bibliográficas, en la minería de datos, en la construcción de tesauros y de ontologías. En general, su objetivo es encontrar semejanzas de significado entre parejas de palabras y semejanzas de significado dentro de agrupaciones de palabras. El análisis de las palabras asociadas (que ha dado lugar a la concepción de múltiples instrumentos, uno de los cuales ha sido utilizado en el presente estudio<sup>2</sup>), ha estado en el centro de la generación de nuevas modalidades de identificación de las estructuras presentes en las bases de datos. Se trata ahora de enfrentar el problema de responder a las necesidades de los investigadores que precisan conocer lo que pasa en sus propios campos para apropiarse de los nuevos resultados generados, para tomar las mejores decisiones y adelantar sus propios estudios.

## MATERIALES Y MÉTODOS

En este documento se presentan, en la primera parte, las características de los datos de la base CAB, la selección, preparación y organización de la información básica; la presentación del método de análisis seleccionado: el método de las palabras asociadas, las condiciones para su aplicación y el tipo de resultados obtenidos; igualmente se presenta en detalle el informe sobre uno de los agregados identificados, *Human Diseases*, que sirve para mostrar las formas estructuradas que adopta la información. En la segunda parte, se interpretan los resultados obtenidos tomando como referencia el mismo agregado. En la tercera parte, se hace uso de las estructuras construidas a partir de la información de base para mostrar explícitamente, bajo la forma de tablas, formas diferentes de evolucionar en el tiempo, y se ponen en relación los documentos que están en la base de la constitución de esas estructuras con los autores, las revistas y las instituciones de los autores. Se señala la necesaria participación de quienes tienen saberes especializados para hacer interpretaciones más contextualizadas. Otros resultados de la aplicación del método, por limitaciones de espacio, no son presentados en el documento. Se termina con algunas observaciones sobre los posibles usos para tomar decisiones en el campo de investigación.

### ESTRUCTURACIÓN DE INFORMACIÓN BIBLIOGRÁFICA POR EL MÉTODO DE LAS PALABRAS ASOCIADAS

Las bases bibliográficas son servicios centralizados que recogen la información de documentos publicados en diferentes áreas del conocimiento, realizan su indexación catalográfica y la integran bajo la forma de fichas bibliográficas que permiten su identificación precisa y orientan la recuperación en diferentes niveles de precisión. Las bases bibliográficas (especializadas o generalistas según recojan documentos en uno, en varios o en todos los campos del conocimiento), por lo general van más allá de la recolección de la información mínima, introduciendo además otros datos que ayudan a precisar las características de los productores y de los contenidos agregando, por ejemplo, las afiliaciones institucionales de los autores, la calificación de los tipos de documentos o asignando descriptores que permiten tener una representación de sus contenidos. La base bibliográfica CAB<sup>3</sup> recoge la información de revistas y de documentos en las áreas de las ciencias de la salud, de la biología y de las ciencias agropecuarias, de bosques y suelos. CAB es la base bibliográfica con mayor representación de la producción latinoamericana, 6,5% de la totalidad de las publicaciones en 11.000 revistas y documentos, lo que señala la importancia regional que tienen estos temas y el nivel de cubrimiento que alcanza en ella (Ricyt, 2001).

<sup>2</sup>Algunos de estos instrumentos, como T-lab y *Réseaux-lu*, están disponibles comercialmente, otros son programas académicos. Para una presentación razonada de algunos de estos instrumentos en el área francófona véase el sitio administrado por C. Lejeune, <http://www.smess.egss.ulg.ac.be/lejeune/logiciels/>

<sup>3</sup>Producida por CABI Publishing, indexa la literatura publicada en el campo de las ciencias agrícolas y ciencias de la vida, cubre cerca de 7.000 revistas y 2.500 libros, así como también reportes técnicos, tesis publicadas, reportes anuales, simposios, libros de texto (*textbooks*) y documentos presentados en eventos. La selección de la literatura es realizada por un equipo de expertos en las disciplinas de interés definidas por CAB, que seleccionan la literatura publicada de 125 países en cerca de 50 idiomas.

1. **La información básica.** Se realizó una extracción de la información bibliográfica de todos los documentos que presentan autores vinculados a instituciones u organizaciones con sede en Colombia durante el período 1997 a 2002. Características de la información: Las fichas bibliográficas tienen los siguientes metadatos:

- i) Título del documento.
- ii) Autores.
- iii) Instituciones de afiliación del primer autor.
- iv) Fuente primaria de los documentos: nombre de la fuente, si se trata de una publicación seriada, año, Vol., No., páginas inicial y final.
- v) Descriptores. Estos descriptores son asignados por documentalistas de manera controlada a partir de un tesoro del que se conoce la estructura y se provee la definición de cada uno de sus términos (lo que permite la interpretación de un conjunto de descriptores asignados a un documento).

2. **Constitución del *corpus* documental.** El *corpus* documental se constituye a partir del conjunto de las fichas bibliográficas de los documentos seleccionados y está formado por los registros ordenados, cada uno con un código numérico que lo identifica y con el conjunto de los descriptores asignados. Un ejemplo de un registro es el siguiente: 716:immunodiagnosis/Human Diseases/ELISA/blood/Chagas' disease/diagnostic techniques /immunofluorescence.

3. **Construcción de los archivos.** Los archivos básicos para realizar los análisis se forman tomando lapsos de tiempo definidos de interés, introduciendo así segmentaciones temporales en el *corpus*. La información disponible en la base bibliográfica se organizó en seis archivos: cada uno cubre un período de un año, de 1997 a 2002. Es posible hacer otras segmentaciones para periodos más amplios.

4. **El método de análisis.** El método de las palabras asociadas es una técnica de representación de la ciencia y la tecnología, que se basa en el análisis de las frecuencias de aparición común de palabras, en nuestro caso descriptores, dentro de un *corpus* de documentos que se considera representativo de uno o varios campos, donde cada uno de los documentos está acompañado de una lista de palabras que sirven para describir su contenido. El programa *Leximappe* permite identificar todos los temas que están relacionados dentro del *corpus*. Cada uno de los temas está formado por un conjunto de descriptores. La conformación de cada uno de ellos depende del grado de asociación entre los descriptores medido por el coeficiente de asociación:

$$E_{ij} = (c_{ij})^2 / c_i \cdot c_j$$

donde  $c_{ij}$  es el número de veces que se encuentran asociados los descriptores  $i$  y  $j$ , y el número de veces que aparecen el descriptor  $i$  y el descriptor  $j$  en el conjunto de documentos respectivamente son  $c_i$  y  $c_j$ .

El método de las palabras asociadas permite establecer los temas que son subconjuntos o agregados de descriptores que se encuentran más fuertemente vinculados entre sí, es decir, con mayor coeficiente de asociación, y que pueden ser interpretados como centros de interés de los investigadores. Si estos agregados tienen muchos descriptores es difícil su interpretación por lo que se fija un número máximo (por ejemplo, diez) y un número mínimo (por ejemplo, tres). Estos parámetros definen a su vez un umbral de saturación que es el valor del primer vínculo rechazado al alcanzar el número máximo de elementos de un agregado, luego de lo cual se procede a formar otro agregado, y así sucesivamente. El análisis de los agregados y de sus vínculos externos (coeficientes de asociación entre descriptores de un agregado con descriptores de otros agregados), permite establecer cuáles nuevos agregados son en realidad prolongaciones de otros que ya se habían formado y cuáles son independientes, en el sentido de no tener vínculos con otros agregados. Así, es posible establecer subredes entre descriptores, formadas por los agregados asociados dentro de la red general, definida por todas las asociaciones entre todos los descriptores. Los agregados aislados o independientes forman subredes temáticas relativamente autónomas.

La centralidad de un agregado se define por la suma de los vínculos externos (entre temáticas o descriptores del mismo agregado con aquellos de otros agregados); y su densidad, por el valor medio de sus vínculos internos. A mayor valor de centralidad de un agregado, hay más importancia acordada por la relación entre las temáticas; y a mayor valor de densidad de un agregado, más desarrollado aparece el tema, representado por el conjunto de descriptores que lo forman. Se puede entonces interpretar que cuando es mayor la centralidad, los investigadores que trabajan en el tema general (determinado por el conjunto de sus temáticas y de sus asociaciones) encuentran también de interés la relación entre estas temáticas y las que se desarrollan en los otros agregados asociados. En el caso de una mayor densidad, el tema general del agregado ha sido abordado y desarrollado por investigadores que lo han llevado a una relativa madurez. La ubicación de cada agregado en un plano de dos dimensiones, cuyos ejes son dados por los valores de densidad y de centralidad, y donde el origen es el valor medio de todos los valores de densidad y de centralidad, permite entonces obtener una imagen de la morfología y de la estructura de la red cognitiva formada por los vínculos entre todos los descriptores. Para la aplicación del método y la generación de resultados se utilizaron los siguientes parámetros:

- Número máximo de descriptores en un agregado igual a diez.
- Número mínimo de descriptores en un agregado igual a tres.
- Frecuencia mínima de un descriptor igual a tres.
- Frecuencia mínima de asociación de descriptores igual a tres.

Aquí no se pretende hacer el estudio exhaustivo de todo el *corpus*, solo se señala cómo la utilización del método permitiría hacer la evaluación de la producción de documentos que circulan en una base bibliográfica especializada. De todas formas los resultados de la aplicación del método permiten hacer preguntas particulares sobre temas específicos, sobre sus contenidos, sobre sus relaciones internas y externas en una generación (en un momento específico), o sobre sus evoluciones (transformaciones sucesi-

vas que van teniendo a lo largo del tiempo). Así, con el fin de establecer relaciones entre agregados a través del tiempo, se formaron archivos por períodos sucesivos. En la Tabla 2 se muestran sus características: la identificación y año al que corresponden, el número de documentos, el número de descriptores utilizados para los documentos en cada uno de los períodos, y el número de los efectivamente utilizados luego de la aplicación de los parámetros. La drástica disminución en el número de descriptores utilizados efectivamente con respecto al total de descriptores, permite tener una apreciación del grado de dispersión temática presente en el conjunto de los documentos. En efecto, no se toman en consideración los descriptores que tienen una frecuencia menor a tres, ni los que se encuentran asociados a otros descriptores con una frecuencia menor a tres.

Archivo	Número de documentos	Número total descriptores	Número de descriptores utilizados luego de la aplicación de los parámetros
De97	101	499	64
De98	321	1363	322
De99	374	1395	314
De00	308	1326	242
De01	348	1519	267
De02	259	1185	194

Tabla 2. Características de los archivos de base.

## RESULTADOS

### PRESENTACIÓN DE LOS RESULTADOS DEL MÉTODO DE LAS PALABRAS ASOCIADAS

La aplicación del método de las palabras asociadas para cada uno de los años del período dio lugar a una serie de informes. La distribución de los agregados para una generación se presenta en un plano de dos dimensiones llamado “diagrama estratégico”, cuyos ejes son los valores de densidad y de centralidad; a cada agregado se le ha atribuido un nombre. En la Figura 1, se muestra el diagrama correspondiente a las relaciones externas del archivo De99; y en la Figura 2, el diagrama correspondiente a las relaciones internas de uno de los agregados de este período.

El interés está en establecer las relaciones entre: a) las temáticas o descriptores en el interior de un agregado o tema general, b) las relaciones al interior de una red determinadas por todos los vínculos entre los descriptores o temáticas en una misma generación, c) las relaciones entre temas en diferentes generaciones. Se entiende por generación, la red derivada de la información de un mismo período en la base bibliográfica. Así, los archivos Det<sub>t</sub> se refieren al período t, y dan lugar a la red de la generación i. Se tiene, entonces, que la red de la primera generación es la generada por todas las temáticas presentes en los documentos del año 1997, la segunda generación las del período 1998, y así sucesivamente. Para facilitar la interpretación de los agregados formados, las relaciones internas entre temáticas y las relaciones externas entre agregados de una misma generación, así como entre agregados de diferentes generaciones, se ha introducido la siguiente notación:  $C_i$  es el j<sup>avo</sup> agregado de la i<sup>ava</sup> generación.



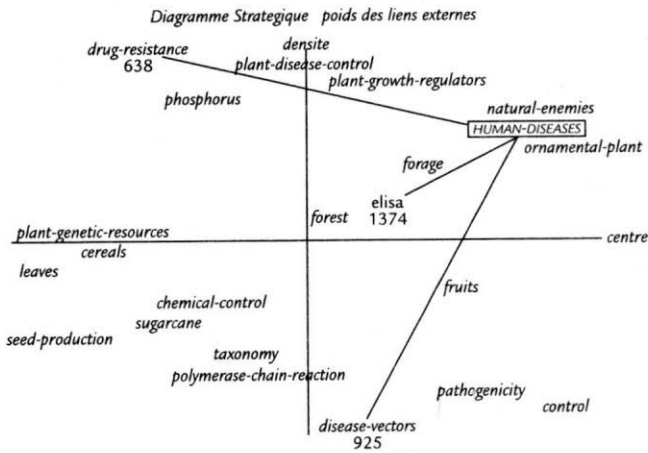


Figura 1. Diagrama estratégico generado a partir de los documentos presentes de coautores que hacen parte de instituciones y entidades en Colombia en la base bibliográfica CAB, integrados en el año 1999. Cálculos OCyT, 2004.

La identificación de las redes temáticas, de los agregados o temas que tienen entre sí valores altos de asociación para un mismo período, permite ubicar los temas que son prolongaciones unos de otros. Por ejemplo, en el diagrama estratégico (Fig. 1), se encuentra que el tema *Human diseases* tiene fuertes relaciones externas con los temas *Elisa*, *Disease vector* y, con valores más bajos, con *Drug resistance*. La inspección del contenido de cada uno de estos temas hace posible la interpretación de los resultados que están en la base de las publicaciones, que han vinculado los temas señalados por medio de los descriptores. La estructura interna del tema *Human Diseases* (Fig. 2.), muestra las relaciones entre los descriptores o temáticas del tema.

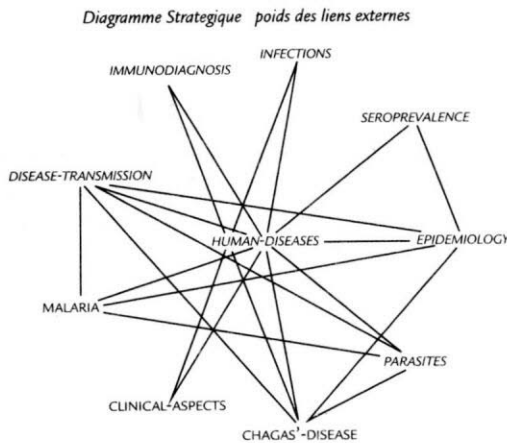


Figura 2. Relaciones internas del tema *Human Diseases* generado a partir de los documentos integrados en la base CAB en 1999, que tienen como coautores a miembros de instituciones o entidades en Colombia. Cálculo OCyT, 2004.

De la estructura interna del tema *Human Diseases*, se dispone también de la siguiente información: el código de identificación del descriptor o temática (número de la izquierda) y su frecuencia (número entre paréntesis a la derecha), en el *corpus* formado por los documentos dentro del período, año 1999. Así, la temática *Human Diseases* está presente en 33 documentos durante el período, y *Epidemiology* está presente en 15. El conjunto de temáticas del tema *Human Diseases* es:

- 6 *Human Diseases* (33)
  - 23 *Epidemiology* (15)
  - 12 *Parasites* (22)
  - 42 *Chagas' Disease* (10)
  - 62 *Clinical Aspects* (8)
  - 118 *Malaria* (6)
  - 26 *Disease transmission* (14)
  - 55 *Immunodiagnosis* (9)
  - 57 *Infections* (9)
  - 159 *Seroprevalence* (5)
- Número de descriptores o temáticas: 10  
 Número de vínculos internos entre las temáticas: 20

A continuación se presentan los valores de los coeficientes de asociación entre los descriptores o temáticas. Por ejemplo, la relación: 23 26 (305), indica que los descriptores *Epidemiology* (23) y *Disease transmission* (26) están vinculados y el valor del coeficiente de esa vinculación es igual a 0,305, que se ha multiplicado por 1.000 (los coeficientes de asociación varían entre 0 y 1). Las relaciones internas entre las temáticas del tema *Human Diseases* son:

23 26 (305)	6 23 (244)	57 62 (222)	23 159 (213)
6 12 (198)	6 42 (194)	6 62 (186)	6 118 (182)
42 55 (178)	6 26 (175)	6 55 (165)	6 57 (165)
12 42 (164)	61 59 (152)	12 118 (121)	12 26 (117)
26 42 (114)	23 42 (107)	26 118 (107)	23 118 (100)

En las relaciones externas del tema *Human Diseases* (Fig. 1), se muestran también el valor de las asociaciones entre los temas *Human Diseases* ( ${}_3C_1$ )<sup>4</sup>, *Elisa* ( ${}_3C_{18}$ ), *Disease vector* ( ${}_3C_{15}$ ) y *Drug resistance* ( ${}_3C_7$ ). A continuación se presentan las relaciones particulares entre las temáticas que tienen un mayor valor de asociación del tema, y que dan lugar a las relaciones entre temas.

- 55 *Immunodiagnosis* > 190 *Immunofluorescence*  ${}_3C_{18}$  (445)
- 55 *Immunodiagnosis* > 83 *Elisa*  ${}_3C_{18}$  (254)
- 42 *Chagas' Disease* > 190 *Immunofluorescence*  ${}_3C_{18}$  (225)
- 55 *Immunodiagnosis* > 132 *Antigens*  ${}_3C_{18}$  (200)

<sup>4</sup>La notación  $C_i$  se utiliza para señalar que el tema o agregado pertenece a la tercera generación,  $i = 3$ , (para 1999) y que corresponde, en esa generación, al tema  $j$ , igual a 18, 15 y 7 en este caso.

26 *Disease transm* > 27 *Disease vector*  ${}_3C_{15}$  (510)

42 *Chagas' Diseases* > 27 *Disease vector*  ${}_3C_{15}$  (179)

12 *Parasites* > 106 *Antimalarials*  ${}_3C_7$  (189)

12 *Parasites* > 76 *Antiprotozoal*  ${}_3C_7$  (162)

12 *Parasites* > 136 *Chloroquine*  ${}_3C_7$  (145)

12 *Parasites* > 63 *Drug resistance*  ${}_3C_7$  (142)

Índice de densidad: 34 Índice de centralidad: 53

## DISCUSIÓN

### SOBRE LA INTERPRETACIÓN DE LOS RESULTADOS

Los resultados de la aplicación del método pueden considerarse como una representación estructurada de las redes de temas que, por períodos, emergen de los vínculos entre documentos debidos a la presencia de temáticas comunes. Ahora bien, se trata de resultados que pueden ser interpretados según cada generación y que deben, además, ser interpretados en conjunto para conocer la estructura del campo, los paradigmas locales que orientan la selección de los problemas de investigación y las formas de enfrentarlos. Los resultados según generaciones, es decir, según los períodos en que se ha fragmentado todo el *corpus* global, se pueden considerar representaciones temporales de la actividad desarrollada por autores cuyos trabajos, estudios e investigaciones dieron lugar a los documentos integrados durante el período en la base bibliográfica. A una escala mayor, las representaciones globales permiten hacer interpretaciones sobre la estructura del campo cubierto por las publicaciones, sobre los paradigmas locales que orientan la selección y sobre las modalidades de enfrentar los problemas de investigación.

Las representaciones globales pueden obtenerse si se consideran las redes temáticas estructuradas identificadas por período, como textos estructurados interpretables. De igual manera, la unión de las redes estructuradas entendida como su yuxtaposición, y la posibilidad de pasar de una a otra por las coincidencias totales o parciales de temáticas comunes, también puede ser interpretada. Esta yuxtaposición da lugar a un hipertexto (Edwards, 1994), que permite la navegación a través de los vínculos definidos y presentes en la unión de estas estructuras, derivadas de la aplicación del método de las palabras asociadas a los conjuntos de documentos de cada generación. La interpretación de los resultados y de sus relaciones pasa ahora por la identificación de cada uno de los elementos, de los tipos de vínculos que se dan entre ellos y de los contextos en que se presentan. Inicialmente la aplicación del método permitió construir los conjuntos de temas para cada generación (Tabla 3).

1997	$C_1 = \{ {}_1C_i : i = 1, 2, \dots, 5 \}$
1998	$C_2 = \{ {}_2C_i : i = 1, 2, \dots, 16 \}$
1999	$C_3 = \{ {}_3C_i : i = 1, 2, \dots, 23 \}$
2000	$C_4 = \{ {}_4C_i : i = 1, 2, \dots, 12 \}$
2001	$C_5 = \{ {}_5C_i : i = 1, 2, \dots, 16 \}$
2002	$C_6 = \{ {}_6C_i : i = 1, 2, \dots, 9 \}$

Tabla 3. Conjuntos de temas identificados por período.

Se pueden identificar varios niveles complementarios de interpretación. Un primer nivel corresponde a la interpretación, para cada generación, de las relaciones internas entre las temáticas que hacen parte de un mismo tema, y de las relaciones externas entre temáticas de un tema que se considera como centro y las de otros temas, éstos últimos pueden ser prolongaciones del tema central. Por otra parte, es posible que existan temas aislados, cuyo valor de centralidad es cercano a cero, lo que indica que se realizan trabajos, estudios e investigaciones en las temáticas interiores señaladas por los descriptores, pero que están cerrados sobre sí mismos sin que tengan relaciones significativas con otras temáticas exteriores (valores bajos de los coeficientes de asociación entre temas). Un segundo nivel considera diferentes *corpus* en el tiempo. En este caso se tiene que hay temas generales que se prolongan en el tiempo, pasando de una generación a otra, transformándose, ya sea por el abandono de algunas temáticas o por la integración de otras. Para una apreciación de las evoluciones de un tema en las distintas generaciones se ha definido el coeficiente de permanencia (Callon *et al.*, 2001) entre dos sucesivas generaciones así:

$$\text{Coeficiente de permanencia } F_{hk} = 2f_{hk} / (f_h + f_k)$$

donde  $f_{hk}$  es el número de temáticas comunes en un tema general en dos generaciones sucesivas,  $f_h$  es el número de temáticas del tema general en la generación antecedente y  $f_k$  es el número de temáticas del tema general en la generación posterior. De manera análoga se puede definir el coeficiente de transformación  $T_{hk}$ , así:

$$\text{Coeficiente de transformación } T_{hk} = 1 F_{hk}$$

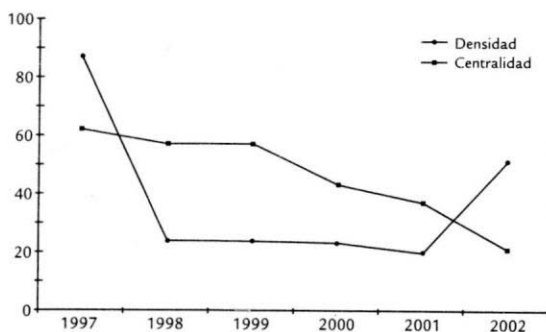


Figura 3. Evolución de los coeficientes estructurales del tema *Human Diseases*.

Se han identificado tres temas generales de los que es posible mostrar esas evoluciones a lo largo del tiempo en que se ha realizado el análisis: *Human Diseases*, *Plant Diseases* y *Biological Control*. Aquí se ha seleccionado el caso de *Human Diseases* del que se muestran, en la Tabla 4, las temáticas que lo forman en cada generación, distinguiendo con caracteres en negrilla las que se conservan cuando se pasa de una generación a la siguiente. También se obtiene una representación de su evolución estructural por la evolución de los coeficientes estructurales de densidad y centralidad (Fig. 3) y del coeficiente de permanencia (Fig. 4). Se realizó la misma presentación para los temas *Plant Diseases* y *Agricultural Entomology*.

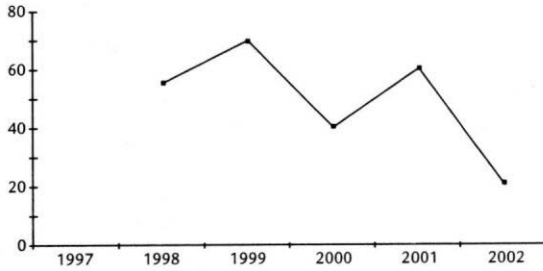


Figura 4. *Human Diseases*. Variación del Índice de permanencia.

El tema general *Plant Diseases* tiene el interés de mostrar que la evolución puede concretarse en el desarrollo de más de un tema particular (Tabla 5), por ejemplo la fragmentación en el año 2001 en dos, uno de ellos evolucionando en la generación siguiente por la integración de temáticas de biotecnología, y el otro conservando la orientación general del tema que tuvo desde sus inicios en 1997.

Generación	1997	1998	1999	2000	2001	2002
	<i>Human disease</i>	<i>Human-Diseases<sup>a</sup></i>	<i>Human-disease</i>	<i>Human-disease</i>	<i>Human-disease</i>	<i>Human-disease</i>
	<i>Epidemiology</i>	<i>Epidemiology</i>	<i>Epidemiology</i>	<i>Diagnosis</i>	<i>Malaria</i>	<i>Malaria</i>
	<i>Parasites</i>	<i>Parasites</i>	<i>Parasites</i>	<i>Case reports</i>	<i>Clinical aspects</i>	<i>Immunization</i>
	<i>Clinical-aspects</i>	<i>Clinical-aspects</i>	<i>Chagas's disease</i>	<i>Clinical-disease</i>	<i>Drug-therapy</i>	<i>Disease models</i>
		<i>Infections</i>	<i>Infections</i>	<i>Infections</i>	<i>Epidemiology</i>	<i>Vaccine development</i>
		<i>Malaria</i>	<i>Clinical-aspects</i>	<i>Drug-therapy</i>	<i>Case-reports</i>	<i>Chloroquine</i>
		<i>Drug-therapy</i>	<i>Malaria</i>	<i>Dwelling</i>	<i>Children</i>	<i>Vaccination</i>
		<i>Cutaneous-leishmaniasis</i>	<i>Disease-transmission</i>	<i>Immunization</i>	<i>Diagnosis</i>	<i>Laboratory-animals</i>
		<i>Case-reports</i>	<i>Immunodiagnosis</i>	<i>Epidemiology</i>	<i>Cytokines</i>	<i>Immune-response</i>
		<i>Disease-transmission</i>	<i>Seroprevalence</i>	<i>Reviews</i>	<i>Paracoccidiodiomycosis</i>	<i>Antigens</i>
(Dens., Cent) <sup>b</sup>	(87, 62)	(24, 57)	(24, 57)	(23, 43)	(20, 37)	(51, 21)
Índice de permanencia <sup>c</sup> F <sub>ik</sub>		0.5	0.7	0.4	0.6	0.2
{C} <sub>j</sub> <sup>d</sup>	<sub>1</sub> C <sub>4</sub>	<sub>2</sub> C <sub>2</sub>	<sub>3</sub> C <sub>1</sub>	<sub>4</sub> C <sub>2</sub>	<sub>5</sub> C <sub>3</sub>	<sub>6</sub> C <sub>2</sub>

Tabla 4. Evolución en el tiempo del tema *HUMAN-DISEASES*.

Convenciones: <sup>a</sup> Se señalan las temáticas que se repiten en dos generaciones sucesivas en letra cursiva; <sup>b</sup> los valores de los coeficientes de densidad y de centralidad (D, C), han sido multiplicados por 100; <sup>c</sup> se ha definido el coeficiente de permanencia (respectivamente el coeficiente de transformación), por la fórmula  $F_{ik} = 2f_{ik} / (f_i + f_k)$  ( $T_{ik} = (1 - 2f_{ik}) / (f_i + f_k)$ ) donde  $f_{ik}$  es el número de temáticas comunes en un tema general en dos generaciones sucesivas,  $f_i$  es el número de temáticas del tema general en la generación antecedente y  $f_k$  es el número de temáticas del tema general en la generación posterior; <sup>d</sup> la notación C<sub>j</sub> se utiliza para señalar que el tema general se considera en la generación  $i$  y que corresponde, en esa generación, al tema  $j$ . Así, por ejemplo <sub>1</sub>C<sub>4</sub> corresponde al tema *HUMAN-DISEASES*, durante la primera generación y allí es el tema número cuatro.

Identificación C <sub>1</sub>	2C <sub>3</sub>	3C <sub>3</sub>	3C <sub>2</sub>	4C <sub>5</sub>	5C <sub>1</sub>	6C <sub>3</sub>
Plant-Diseases	Plant-Pathogenic-Fungi	Plant-Diseases	Plant-Disease-	Plant-Diseases Control	Plant-Disease-	Genetic-Markers
Plant-Diseases	Plant-Pathogens	Plant-Pathogens	Plant-Pathogenic-Fungi		Fungal-Antagonists	Microsatellites
Plant-Diseases	Plant-Diseases	Plant-Pathogenic-Fungi	Antagonism	Plant-Pathogenic-Fungi	Cultural-Control	Linkage-Groups
Plant-Diseases	Plant-Pathology	Plant-Pathology	Antagonists		Biologica-Control	Genetic Diversity
Plant-Diseases	Fodder-Plants	Plant-Disease-Control	Biological-Control	Plant-Disease-Control	Tomatoes	Loci
Plant-Diseases	Grasses	Plant-Pathogenic-Bacteria	Plant-Diseases		Antagonism	Genetic-Polymorphism
Plant-Diseases	Disease-Resistance	Disease-Resistance	Chrysanthemums	Passion-Fruits	Fungal-Diseases	Genotypes
Plant-Diseases		Resistance	Plant-Pathogens		Chemical-Control	Genes
Plant-Diseases		Avocados	Biological-Control-Agents	Cultivars	Plant-Diseases	Disease-Resistance
Plant-Diseases		Crop Damage	Inoculation	Chemical-Control	Biological-	Plant-Diseases
				Control-Agents		
				Pathogenicity		
				Germplasm		
				Plant-Pests		
(Dens., Cent.)	(88, 0)	(53, 8)	(73, 14)	(11, 48)	(54, 38)	(27, 36)
Índice de permanencia		0.6	0.3	0.3	0.3	0.1

Identificación C <sub>1</sub>	5C <sub>4</sub>	6C <sub>5</sub>
Plant-Diseases	Plant-Pathogens	Plant-Pathogens
Plant-Diseases	Plant-Pathogenic-Fungi	Plant-Pathogenic-Fungi
Plant-Diseases	Disease-Resistance	Fungal-Diseases
Plant-Diseases	Plant-Pathogenic-Bacteria	Plant-Pathogenic-Bacteria
Plant-Diseases	Fungicides	Methodology
Plant-Diseases	Items	
Plant-Diseases	Seeds	
Plant-Diseases	Cultivars	
Plant-Diseases	Fruits	
Plant-Diseases	Leaves	
(Dens., Cent.)	(15, 49)	(27, 26)
Índice de permanencia	0.2	0.4

Tabla 5. Evolución en el tiempo del tema PLANT DISEASES.

El tema general *Agricultural Entomology*, permite también mostrar el interés en extender la noción de generación a períodos más amplios de los que se consideraron inicialmente. En efecto, tal como se encuentra a partir de los datos consignados en la Tabla 6, para el tema *Agricultural Entomology* no hay presencia de él en los *corpus* documentales de los años 2000 y 2002, lo que lleva a la necesidad de explorar la posibilidad de constituir *corpus* que agrupen períodos más amplios para seguir la evolución de un tema. Es posible, también, explorar la posibilidad de que haya un agotamiento del tema general.

Un tercer nivel de interpretación es posible cuando se asocian a los temas los documentos, los autores y las instituciones. En la Tabla 7 se presenta para el tema *Human Diseases*, y en la Tabla 8 para *Plant Diseases*, las contribuciones a esos temas de los documentos, de sus autores y de las instituciones a las que pertenecen en el momento

de su publicación. La asociación entre el número de los documentos y las instituciones de los primeros autores establece las contribuciones y los compromisos institucionales con el tema, tanto por cada año como para todo el período considerado. Por ejemplo, en el tema *Human Diseases* la mayor parte de los documentos provienen de centros e institutos de investigación: Centro Internacional de Entrenamiento e Investigaciones Médicas, CIDEIM (21 documentos, 72 autores o coautores); Instituto Nacional de Salud, INS (20, 72); Corporación de Investigaciones Biológicas, CIB (16, 77); Instituto Colombiano de Medicina Tropical, ICMT (7, 33); o de universidades del Estado: Universidad de Antioquia, UA (21, 65); Universidad Nacional de Colombia, UN (14, 56); Universidad del Valle (9, 35); y Universidad Industrial de Santander, UIS (5, 30).

Generación	1997	1998	1999	2000	2001	2002
Agricultural-Entomology	Agricultural Entomology	Agricultural Entomology	Natural-Enemies		Entomogenous-Fungi	
Agricultural-Entomology	Plant-Pests	Pest-Control	Parasitoids		Entomopathogens	
Agricultural Entomology	Conidia	Control	Agricultural Entomology		Conidia	
Agricultural Entomology	Mortality Enemies	Natural-Insects	Beneficial-		Predators	
Agricultural Entomology	Formulations	Biology	Hosts		Plant-Pests	
Agricultural Entomology	Geographical-Distribution	Evaluation	Longevity		Mortality	
Agricultural Entomology	Pathogenicity	Integrated-Pest-Management	Predators		Insect-Pests	
Agricultural Entomology	Toxicity	Damage	Entomopathogens		Pathogenicity	
Agricultural Entomology	Control Stimulant Plants	Hosts Plant-Pests	Insect-Control		Enzymes Natural-Enemies	
(Dens., Cent.)	(21, 113)	(25, 30)	(35, 73)		(44, 28)	
Índice de Permanencia		0.2	0.3		0.2	
{C <sub>s</sub> }	,C <sub>2</sub>	,C <sub>4</sub>	,C <sub>3</sub>		,C <sub>2</sub>	

Tabla 6. Evolución del tema AGRICULTURAL-ENTOMOLOGY en el tiempo.

En el tema *Plant Diseases* la presencia de la Universidad Nacional de Colombia es hegemónica con 76 documentos y 157 autores o coautores, seguida a distancia por los centros de investigación: Centro Internacional de Agricultura Tropical, CIAT (61, 201); Corporación Colombiana de Investigación Agropecuaria, CORPOICA (20, 44); y el Centro Nacional de Investigaciones de Café, CENICAFE (9, 31). Un resultado derivado de este tipo de asociaciones es el conocimiento de las revistas que son más utilizadas para la publicación de los documentos.

La identificación de las asociaciones de temáticas dentro de temas y de sus evoluciones en el tiempo permite asimismo, identificar los actores y las redes socio cognitivas en que se mueven, sus movimientos y actividades dentro de instituciones, el reconocimiento de sus aportes, y su participación en la generación de nuevos conceptos, en la integración de nuevas técnicas, y en el desarrollo de nuevos procedimientos que atraen a otros interesados, y que luego de un cierto tiempo devienen teorías aceptadas o técnicas rutinizadas. Se abre, entonces, la posibilidad de interpretar las redes temáticas como redes socio cognitivas, de actores alrededor de referencias compartidas, desde la perspectiva

de la existencia o de la emergencia de paradigmas locales en temas o dominios particulares, que pueden considerarse desde dos ópticas complementarias:

- a) Como la presencia de comunidades locales que participan de las mismas referencias teóricas para desarrollar sus trabajos y que, entonces, pueden poner a disposición los nuevos resultados logrados o disponer de las contribuciones particulares de otros para beneficio propio y de los otros miembros.
- b) Como la existencia de nociones, conceptos y teorías de referencia que sirven de orientación para las actividades de los investigadores, que desde diferentes lugares institucionales contribuyen a la emergencia de una comunidad.

Esta interpretación hace posible también considerar el espacio de las decisiones políticas orientadas por el conocimiento de las redes y de los actores, de las referencias compartidas y de los saberes disponibles. Estos niveles de interpretación deberían complementarse con los aportes y los juicios calificados de los especialistas que tienen saberes especializados y experiencias en el campo. Este último nivel va más allá de las posibilidades de quienes realizaron este estudio.

#### ORIENTACIONES PARA LA UTILIZACIÓN DE LA INFORMACIÓN ESTRUCTURADA

Es posible sacar provecho a los resultados obtenidos si se tiene en cuenta la teoría utilizada, donde se vinculan los actores, sus producciones y que se concretan en los documentos. En primer término, es posible mostrar las evoluciones de los temas por abandono de ciertas temáticas y la inserción de otras, los actores que participan en ellas, y las instituciones a las que a éstos pertenecen. En segundo lugar, es posible la generación de estrategias de los propios investigadores a partir de las interpretaciones de los resultados: opciones de investigación, constitución de bases de datos particulares sobre temas específicos, constitución de alianzas y colaboración entre grupos de investigación. A continuación se muestran las evoluciones y a manera de tablas, se señalan las relaciones entre datos heterogéneos: autores, temáticas, instituciones, evoluciones, técnicas. Las decisiones tomadas con base en este tipo de información estructurada y las interpretaciones que de ellas se desprenden son, por supuesto, asunto de los propios investigadores.

Generación {C <sub>i</sub> }	Instituciones	Documentos	Autores	Revistas
1C <sub>4</sub>	Instituto Nacional de Salud (INS)	4	19	Biomédica
	Secretaría de Salud del Guainía	1	2	Biomédica
	Universidad Industrial de Santander (UIS)	1	7	Biomédica
2C <sub>2</sub>	Centro Internacional de Entrenamiento e Investigaciones Médicas (CIDEIM)	9	30	Acta Tropical
	Instituto Colombiano de Medicina Tropical ICMT	4	19	Lancet British Edition American Journal of Tropical Medicine and Hygiene Biomédica Memórias do Instituto Oswaldo Cruz American Journal of Tropical Medicine and Hygiene
	Corporación para Investigaciones Biológicas (CIB)	3	20	Revista do Instituto de Medicina Tropical de Sao Paulo



Generación {C <sub>i</sub> }	Instituciones	Documentos	Autores	Revistas
2C <sub>2</sub>	Corporación para Investigaciones Biológicas (CIB)	3	20	<i>Memorias do Instituto Oswaldo Cruz</i> <i>Biomédica</i> <i>Journal of Clinical Microbiology</i>
	Instituto Nacional de Salud (INS)	3	14	<i>Clinical and Diagnostic Laboratory Immunology</i> <i>Revista do Instituto de Medicina Tropical de Sao Paulo</i> <i>Biomédica</i>
	Universidad del Valle (UV)	2	10	<i>Annals of Internal Medicine</i> <i>Annals of Tropical Medicine and Parasitology</i>
	Ministerio de Salud	1	4	<i>Biomédica</i> <i>Cadernos de Saude Publica</i>
	Fundación Clínica Shaio	1	10	<i>Revista de Patología Tropical</i>
	Secretaría de Salud del Guainía	1	2	<i>Biomédica</i>
	Universidad de Antioquia (UA)	1	6	<i>American Journal of Tropical Medicine and Hygiene</i>
	Universidad de Caldas (UC)	1	2	<i>Archives of Dermatology</i>
	Universidad Industrial de Santander (UIS)	1	5	<i>Biomédica</i>
	Universidad Nacional de Colombia (UN)	1	9	<i>Journal of Peptide Research</i>
3C <sub>1</sub>	Centro Internacional de Entrenamiento e Investigaciones Medicas (CIDEIM)	6	21	<i>Memorias do Instituto Oswaldo Cruz</i>
	Universidad de Antioquia (UA)	5	17	<i>Biomédica</i>
	Corporación para Investigaciones Biológicas (CIB)	3	12	<i>American Journal of Tropical Medicine and Hygiene</i> <i>Journal of the American Academy of Dermatology</i> <i>Memorias do Instituto Oswaldo Cruz</i>
	Instituto Nacional de Salud (INS)	3	11	<i>Biomédica</i> <i>Revista Iberoamericana de Micología</i> <i>American Journal of Tropical Medicine and Hygiene</i> <i>Revista Iberoamericana de Micología</i> <i>Journal of Clinical Microbiology</i>
	Universidad Nacional de Colombia (UN)	3	14	<i>Clinical Infectious Diseases</i> <i>Biomédica</i> <i>Parasitology Research</i>
	Universidad de los Andes (UA)	2	3	<i>Biomédica</i> <i>Memorias do Instituto Oswaldo Cruz</i>
	Consortio de Investigaciones Bioclínicas (CIBIC)	1	7	<i>Medicina</i>
	Academia Nacional de Medicina	1	8	<i>Clinical Infectious Diseases</i> <i>Clinical and Diagnostic Laboratory Immunology</i>
	Hospital Universitario San Vicente de Paul	1	4	<i>Clinical Transplantation</i>
	Instituto Colombiano Agropecuario (ICA)	1	4	<i>Revista de Saude Publica</i>
	Instituto Colombiano de Medicina Tropical ICMT	1	6	<i>Revista Panamericana de Salud Pública</i>
	Instituto de Salud del Pacífico (INSALPA)	1	3	<i>Social Science and Medicine</i>
	Instituto de Salud en el Trópico	1	10	<i>Revista do Instituto de Medicina Tropical de Sao Paulo</i>

34 Artículo - Estructuración de la información bibliográfica en la base CAB de autores vinculados a instituciones colombianas, 1997 a 2002. Charum, Usgame.

Generación { <sub>i</sub> , <sub>j</sub> }	Instituciones	Documentos	Autores	Revistas
3C <sub>1</sub>	Secretaría de Salud del Amazonas	1	7	Biomédica
	Universidad de Caldas (UC)	1	7	Revista Iberoamericana de Micología
	Universidad del Tolima (UT)	1	4	Acta Tropical
	Universidad Industrial de Santander (UIS)	1	9	Biomédica
4C <sub>2</sub>	Universidad de Antioquia (UA)	7	19	Biomédica
	Instituto Nacional de Salud (INS)	3	7	Perspectivas en Nutrición Humana
	Corporación para Investigaciones Biológicas (CIB)	2	6	<i>Transactions of the Royal Society of Tropical Medicine and Hygiene</i>
	Universidad Industrial de Santander (UIS)	2	9	<i>Clinical and Diagnostic Laboratory Immunology</i> <i>International Journal of Epidemiology</i> Revista Panamericana de Salud Pública Biomédica <i>Revista do Instituto de Medicina Tropical de Sao Paulo</i>
	Fundación para la Educación Superior (FES)	1	8	<i>Medical Mycology</i> <i>Transactions of the Royal Society of Tropical Medicine and Hygiene</i>
	Instituto Colombiano de Medicina Tropical (ICMT)	1	4	Investigación Clínica <i>American Journal of Tropical Medicine and Hygiene</i>
	Instituto de Salud del Pacífico (INSALPA)	1	3	Biomédica <i>Transactions of the Royal Society of Tropical Medicine and Hygiene</i>
	Universidad de los Andes (UA)	1	5	<i>Memorias do Instituto Oswaldo Cruz</i>
	5C <sub>3</sub>	Universidad Nacional de Colombia (UN)	9	30
Instituto Nacional de Salud (INS)		6	19	<i>Journal of Clinical Microbiology</i>
Corporación para Investigaciones Biológicas (CIB)		5	26	Biomédica
Universidad de Antioquia (UA)		4	15	<i>British Journal of Dermatology</i> <i>Environmental Health Perspectives</i> <i>International Journal for Parasitology</i> <i>Experimental Parasitology</i> Biomédica
Universidad del Valle (UV)		4	8	<i>Medical Mycology</i> <i>Mycopathology</i> <i>Infection and Immunity</i>
Centro Internacional de Entrenamiento e Investigaciones Médicas (CIDEIM)		1	4	<i>Seminars in Arthritis and Rheumatism</i> <i>Cadernos de Saude Publica</i> Revista Panamericana de Salud Pública <i>American Journal of Chinese Medicine</i> Revista Colombiana de Ciencias Pecuarias
Fundación Canguro		1	2	<i>Transactions of the Royal Society of Tropical Medicine and Hygiene</i>

Generación {C <sub>j</sub> }	Instituciones	Documentos	Autores	Revistas
5C <sub>3</sub>	Fundación Canguro	1	2	<i>Memorias do Instituto Oswaldo Cruz</i> <i>Health Policy and Planning</i> <i>Biomédica</i>
	Fundación para la Educación Superior (FES)	1	5	<i>Molecular Immunology</i> <i>American Journal of Tropical Medicine and Hygiene</i> <i>Acta Pediátrica</i> <i>Cardiology</i>
	Hospital Eduardo Arredondo	1	6	<i>Biomédica</i>
	Hospital Meissen	1	4	<i>Biomédica</i>
	Instituto Colombiano de Medicina Tropical (ICMT)	1	4	<i>Biomédica</i>
	Pontificia Universidad Bolivariana	1	6	<i>Epidemiology and Infection</i>
	Pontificia Universidad Javeriana (PUJ)	1	1	<i>International Journal of Tuberculosis and Lung Disease</i>
	Universidad Autónoma de Bucaramanga	1	1	<i>International Journal of Epidemiology</i>
	Universidad Surcolombiana	1	5	<i>Revista Panamericana de Salud Pública</i>
6C <sub>2</sub>	Fundación Instituto de Inmunología de Colombia (FIDIC)	7	39	<i>Journal of Peptide Research</i>
	Centro Internacional de Entrenamiento e Investigaciones Médicas (CIDEIM)	4	12	<i>Life Sciences</i> <i>Vaccine</i> <i>FEBS Letters</i> <i>Infection and Immunity</i> <i>Experimental Parasitology</i> <i>Memorias do Instituto Oswaldo Cruz</i>
	Universidad de Antioquia (UA)	4	8	<i>Infection and Immunity</i> <i>Revista Panamericana de Salud Pública</i>
	Universidad del Valle (UV)	3	17	<i>Revista Colombiana de Ciencias Pecuarias</i> <i>Clinical Biochemistry</i> <i>Memorias do Instituto Oswaldo Cruz</i> <i>Parasite Immunology</i>
	Corporación para Investigaciones Biológicas (CIB)	2	6	<i>Journal of Parasitology</i> <i>International Journal for Parasitology</i> <i>Journal of Comparative Pathology</i>
	Instituto Nacional de Salud (INS)	1	2	<i>Veterinary Parasitology</i> <i>Biomédica</i>
	Instituto de Salud en el Trópico	1	4	<i>Revista do Instituto de Medicina Tropical de Sao Paulo</i>
	Universidad Nacional de Colombia (UN)	1	3	<i>Journal of Antimicrobial Chemotherapy</i>

Tabla 7. Evolución de los vínculos de instituciones, autores y revistas con el tema *Human Diseases*.

Generación {i,C <sub>j</sub> }	Instituciones	Documentos	Autores	Revistas
1C <sub>3</sub>	Universidad Nacional de Colombia (UN)	10	11	Informe de Investigaciones Smurfit Cartón de Colombia Investigación Forestal Agronomía Colombiana ASCOLFI Informa
	Centro Internacional de Agricultura Tropical (CIAT)	2	5	<i>Canadian Journal of Plant Pathology</i> Pasturas Tropicales
	Corporación Colombiana de Investigación Agropecuaria (CORPOICA)	1	3	ASCOLFI Informa
	Corporación Nacional de Investigación y Fomento Forestal (CONIF)	1	2	Serie Técnica Corporación Nacional de Investigación y Fomento Forestal
2C <sub>5</sub>	Universidad Nacional de Colombia (UN)	19	35	Fitopatología Colombiana Revista Facultad Nacional de Agronomía Medellín Agronomía Colombiana ASCOLFI Informa
	Centro Internacional de Agricultura Tropical (CIAT)	7	12	<i>Euphytica</i> Manejo Integrado de Plagas <i>Crop Science</i> Fitopatología Colombiana ASCOLFI Informa
	Corporación Colombiana de Investigación Agropecuaria (CORPOICA)	5	12	<i>Plant Disease</i> ASCOLFI Informa Infomusa Fitopatología Colombiana
	Universidad de Caldas (UC)	3	5	Fitopatología Colombiana
	Universidad Pedagógica y Tecnológica de Colombia (UPTC)	3	8	Fitopatología Colombiana
	Centro de Investigación de la Caña de Azúcar de Colombia (CENICANA)	2	7	Fitopatología Colombiana Carta Trimestral CENICANA
	Instituto Colombiano Agropecuario (ICA)	2	7	Manejo Integrado de Plagas ASCOLFI Informa
	Universidad Católica de Manizales	1	3	Fitopatología Colombiana
	Universidad Santiago de Cali	1	4	Fitopatología Colombiana
	3C <sub>2</sub>	Centro Internacional de Agricultura Tropical (CIAT)	16	52
Centro Internacional de Agricultura Tropical (CIAT)		16	52	<i>Agricultural Biotechnology Policy Seminars</i>
Corporación Colombiana de Investigación Agropecuaria (CORPOICA)		8	19	ASCOLFI Informa Revista Facultad Nacional de Agronomía, Medellín Revista Colombiana de Entomología Fitopatología Colombiana

Generación {C <sub>i</sub> }	Instituciones	Documentos	Autores	Revistas
3C <sub>2</sub>	Centro Nacional de Investigaciones de Café (CENICAFE)	3	9	Boletín Técnico Cenicafe Revista Colombiana de Entomología
	Instituto Colombiano Agropecuario (ICA)	3	6	Fitopatología Colombiana
	Pontificia Universidad Javeriana (PUJ)	3	7	Agronomía Colombiana ASCOLFI Informa Revista Colombiana de Entomología
	Universidad de Caldas (UC)	3	10	Acta Horticulturae Fitopatología Colombiana
	Centro de Investigación de la Caña de Azúcar de Colombia (CENICANA)	2	7	Serie Divulgativa CENICANA <i>Proceedings of the XXIII ISSCT Congress, New Delhi, India, 22-26 February, 1999. Volume 2. 1999, 390-397; 18 ref.</i>
	Prohosa Ltda.	2	3	Fitopatología Colombiana
	Universidad Jorge Tadeo Lozano	2	5	Acta Horticulturae
	Asociación Colombiana de Exportadores de Flores (ASOCOLFLORES)	1	2	Acta Horticulturae
	BASF Colombia	1	4	Acta Horticulturae
	Corporación Biotec	1	4	ASCOLFI Informa
	Corporación CorpoGen	1	5	Acta Horticulturae
	Corporación Nacional de Investigación y Fomento Forestal (CONIF)	1	2	Boletín de Protección Forestal
	Corporación para Investigaciones Biológicas (CIB)	1	3	Revista Facultad Nacional de Agronomía Medellín
	Flores Mocari	1	4	Acta Horticulturae
	Fundación Nova Hylaea	1	1	Entomólogo
	Universidad de Córdoba	1	3	ASCOLFI Informa
	Universidad del Cauca	1	3	Fitopatología Colombiana
Universidad Nacional de Colombia (UN)	23	54	Fitopatología Colombiana Acta Horticulturae Revista Facultad Nacional de Agronomía, Medellín Revista Colombiana de Entomología ASCOLFI Informa Acta Agronómica	
Universidad Pedagógica y Tecnológica de Colombia (UPTC)	2	10	Fitopatología Colombiana ASCOLFI Informa	
Universidad Santiago de Cali	1	3	Revista Colombiana de Entomología	
4C <sub>5</sub>	Centro Internacional de Agricultura Tropical (CIAT)	10	32	Acta Horticulturae <i>Canadian Journal of Plant Pathology</i> <i>Theoretical and Applied Genetics</i> Fitopatología Colombiana <i>Plant Disease</i> <i>Crop Science</i> <i>Field Crops Research</i> <i>Plant Pathology</i>
	Universidad Nacional de Colombia (UN)	7	18	Fitopatología Colombiana Revista Colombiana de Entomología ASCOLFI Informa O Biológico
	Corporación Colombiana de Investigación Agropecuaria (CORPOICA)	5	10	ASCOLFI Informa Fitopatología Colombiana

Generación {C <sub>i</sub> }	Instituciones	Documentos	Autores	Revistas
4C <sub>5</sub>	Universidad Pedagógica y Tecnológica de Colombia (UPTC)	1	5	Fitopatología Colombiana
5C <sub>1</sub>	Universidad Nacional de Colombia (UN)	11	24	Revista Facultad Nacional de Agronomía Medellín Revista Colombiana de Entomología Fitopatología Colombiana Revista de Biología Tropical
	Corporación Colombiana de Investigación Agropecuaria (CORPOICA)	10	28	IOBC WPRS Working Group "Biological Control of Fungal and Bacterial Plant Pathogens". Proceedings of the sixth meeting, Biocontrol Agents: Mode of Action and Interaction with Other Means of Control, Sevilla, Spain, November 30 December 3, 2000 Revista Colombiana de Entomología Fitopatología Colombiana Infomusa Molecular Ecology Notes
	Centro Internacional de Agricultura Tropical (CIAT)	5	20	Centro Internacional de Agricultura Tropical CIAT Fitopatología Colombiana Acta Agronómica Canadian Journal of Microbiology
	Centro Nacional de Investigaciones de Café (CENICAFE)	3	10	Manejo Integrado de Plagas
	Pontificia Universidad Javeriana (PUJ)	3	8	Fitopatología Colombiana ASCOLFI Informa
	Corporación para Investigaciones Biológicas (CIB)	2	4	Memorias do Instituto Oswaldo Cruz Mycopathologia
	BAYER S.A.	1	3	ASCOLFI Informa
	Federación Nacional de Cafeteros de Colombia	1	4	Revista Colombiana de Entomología
	Institut de Recherche pour le Développement (IRD)	1	10	Journal of Chemical Ecology
	Universidad de Caldas (UC)	1	2	Infomusa
Universidad Jorge Tadeo Lozano	1	1	Manejo Integrado de Plagas	
6C <sub>3</sub>	Centro Internacional de Agricultura Tropical (CIAT)	10	37	Plant Pathology Euphytica Molecular Ecology Notes Manejo Integrado de Plagas y Agroecología Crop Science Phytopathology Theoretical and Applied Genetics Genetic Resources and Crop Evolution
	Centro Nacional de Investigaciones de Café (CENICAFE)	3	12	Mycological Research Theoretical and Applied Genetics Genetic Resources and Crop Evolution
	Universidad de Antioquia (UA)	1	5	Journal of Clinical Virology
5C <sub>4</sub>	Centro Internacional de Agricultura Tropical (CIAT)	7	31	Novon Theoretical and Applied Genetics Euphytica Fitopatología Colombiana

Generación {C <sub>j</sub> }	Instituciones	Documentos	Autores	Revistas
5C <sub>4</sub>	Universidad Nacional de Colombia (UN)	6	15	Revista Facultad Nacional de Agronomía Medellín Fitopatología Colombiana Fitoterapia
	Corporación Colombiana de Investigación Agropecuaria (CORPOICA)	5	9	AOBC WPRS Working Group "Biological Control of Fungal and Bacterial Plant Pathogens". Proceedings of the sixth meeting, Biocontrol Agents: Mode of Action and Interaction with Other Means of Control, Sevilla, Spain, November 30 December 3, 2000. <i>Bul</i>
	Corporación Colombiana de Investigación Agropecuaria (CORPOICA)	5	9	ASCOLFI Informa Fitopatología Colombiana
	Instituto Colombiano Agropecuario (ICA)	2	6	Fitopatología Fitopatología Colombiana
	Centro de Investigación de la Caña de Azúcar de Colombia (CENICANA)	2	5	Fitopatología Colombiana <i>International Society of Sugar Cane Technologists</i>
	Pontificia Universidad Javeriana (PUJ)	2	7	Fitopatología Colombiana ASCOLFI Informa
	BAYER S.A.	1	3	ASCOLFI Informa
	Instituto Nacional de Salud (INS)	1	3	Revista Iberoamericana de Micología
	Universidad de Caldas (UC)	1	2	Infomusa
6C <sub>5</sub>	Centro Internacional de Agricultura Tropical (CIAT)	4	12	<i>Plant Pathology</i> <i>Phytopathology</i> <i>Plant Disease</i> <i>American Journal of Potato Research</i>
	Corporación Colombiana de Investigación Agropecuaria (CORPOICA)	1	3	ASCOLFI Informa
	Syngenta	1	3	<i>Bulletin OILB SROP</i>

Tabla 8. Evolución de los vínculos de instituciones, autores y revistas con el tema *PLANT DISEASES*.

## CONCLUSIONES

La aplicación del método de las palabras asociadas al *corpus* de información constituido por los documentos presentes en la base bibliográfica CAB de autores vinculados a instituciones con domicilio en Colombia, permitió la identificación de grandes temas desarrollados en el país en el campo de las ciencias agrícolas y ciencias de la vida durante el período 1997-2002, construir representaciones de las dinámicas de investigación, de los actores y sus contribuciones, y de los temas privilegiados por las instituciones u organizaciones que los albergan.

Se abre la posibilidad de generar alianzas entre quienes han avanzado en temas semejantes o complementarios para avanzar de manera conjunta, de establecer colaboraciones entre quienes han desarrollado técnicas o generado nuevos conocimientos que pueden ponerse a disposición de otros. Se ha pasado así, a establecer las relaciones entre datos heterogéneos: instituciones y temas de investigación, las revistas y la evolución y transformación de los temas a través del tiempo, y la integración de nuevas técnicas en la investigación.

## AGRADECIMIENTOS

El estudio fuente de este artículo recibió el apoyo del Observatorio Colombiano de Ciencia y Tecnología, y se desarrolló en el marco de los trabajos del grupo de investigación en bibliometría de esa institución.

## BIBLIOGRAFÍA

CALLON M, COURTIAL JP, TURNER B. La Méthode Leximappe: Un outil pour l'analyse stratégique du développement scientifique et technique. En: Vinck D, Editor. Gestion de la recherche. Nouveaux problèmes, nouvelles méthodes. Bruselas: de Boeck Wesmel; 2001. p. 207-277.

CHARUM J. El sistema de información, la normalización de la información y la estandarización de los objetos tecnológicos. Ocyt. 2004. Documento de Trabajo. No. 28.

DE SOLLA PRICE DJ. Networks of Scientific Papers. Scientometrics. 1965;149:510-515.

EDWARDS P. Hyper Text and Hypertension: Poststructuralist Critical Theory, Social Studies of Science and Software. Soc Stud Sci. 1994;24:229-278.

LEJEUNE C. Représentations des réseaux de mots associés. 7<sup>me</sup> Journées internationales d'analyse statistique des données textuelles. Louvain La Neuve; 2004. Marzo 10-12.

LEYDESDORFF L. Various Methods for the Mapping of Science. Scientometrics. 1987;11(5 6):295-324.

POLANCO X. Aux sources de la bibliométrie. En: Noyer J. M. (Dir), Les sciences de l'information: bliométrie, scientométrie, infométrie, 13-78. Presses Universitaires de Rennes. Rennes. 1995.

SMALL H, SWEENEY E. Clustering the Science Citation Index Using Co Citations. I. Comparison Methods Scientometrics. 1985;7(3-6):391-409.

RICYT. El estado de la ciencia. Principales indicadores de ciencia y tecnología iberoamericanos/interamericanos. Buenos Aires: Ricyt. 2001.