





**INSTITUTO POLITÉCNICO NACIONAL  
SECRETARÍA DE INVESTIGACIÓN Y POSGRADO**

**CENTRO DE INVESTIGACIONES ECONÓMICAS,  
ADMINISTRATIVAS Y SOCIALES**



**T E S I S**

**ANÁLISIS CIENCIOMÉTRICO EN EL DESARROLLO DE TEMAS  
DE TESIS DE MAESTRÍA EN EL PERIODO 1996-2007. CASO DE  
ESTUDIO: “CENTRO DE INVESTIGACIÓN EN COMPUTACIÓN  
DEL INSTITUTO POLITÉCNICO NACIONAL”**

**Que para obtener el grado de:**

**Maestra en Política y Gestión del Cambio Tecnológico.**

**P R E S E N T A:**

**MARIBEL VAZQUEZ GALLO**

**Director de Tesis:**

**Dra. María del Pilar Longar Blanco  
Dr. Adolfo Guzmán Arenas**

**MÉXICO, D.F.**

**Agosto de 2010**



SIP-14-BIS

# INSTITUTO POLITÉCNICO NACIONAL

## SECRETARÍA DE INVESTIGACIÓN Y POSGRADO

### ACTA DE REVISIÓN DE TESIS

En la Ciudad de MÉXICO siendo las 12:00 horas del día 15 del mes de JUNIO del 2010 se reunieron los miembros de la Comisión Revisora de Tesis, designada por el Colegio de Profesores de Estudios de Posgrado e Investigación de CIECAS para examinar la tesis titulada:

"ANÁLISIS CIENCIOMÉTRICO EN EL DESARROLLO DE TEMAS DE TESIS DE MAESTRÍA EN EL PERÍODO 1996-2007. CASO DE ESTUDIO: CENTRO DE INVESTIGACIÓN EN COMPUTACIÓN DEL INSTITUTO POLITÉCNICO NACIONAL"

Presentada por el alumno:

VÁZQUEZ  
Apellido paterno

GALLO  
Apellido materno

MARIBEL  
Nombre(s)

Con registro: 

B	0	6	1	1	1	8
---	---	---	---	---	---	---

aspirante de:

MAESTRÍA EN POLÍTICA Y GESTIÓN DEL CAMBIO TECNOLÓGICO

Después de intercambiar opiniones los miembros de la Comisión manifestaron **SU APROBACIÓN DE LA TESIS**, en virtud de que satisface los requisitos señalados por las disposiciones reglamentarias vigentes.

### LA COMISIÓN REVISORA

Directores de tesis

DRA. MARÍA DEL PÍLAR LONGAR BLANCO

DR. ADOLFO GUZMÁN ARENAS

DRA. HORTENSIA GÓMEZ VIQUEZ

DR. GILBERTO LORENZO MARTÍNEZ LUNA

DR. HUMBERTO MERRIT TAPIA

PRESIDENTE DEL COLEGIO DE PROFESORES



DR. HUMBERTO MONTEIL GONZÁLEZ  
SECRETARÍA DE EDUCACIÓN PÚBLICA  
INSTITUTO POLITÉCNICO NACIONAL  
CENTRO DE INVESTIGACIONES  
ECONÓMICAS ADMINISTRATIVAS  
Y SOCIALES



**INSTITUTO POLITÉCNICO NACIONAL**  
**SECRETARÍA DE INVESTIGACIÓN Y POSGRADO**

*CARTA CESIÓN DE DERECHOS*

En la Ciudad de México, D.F el día 23 de Junio del 2010, la que suscribe **MARIBEL VAZQUEZ GALLO** alumna del Programa de Maestría en POLITICA Y GESTION DEL CAMBIO TECNOLOGICO, con número de registro B061118, adscrita al CIECAS, manifiesta que es autora intelectual del presente trabajo de Tesis bajo la dirección de la Dra. María del Pilar Longar Blanco y del Dr. Adolfo Guzmán Arenas y cede los derechos del trabajo intitulado “ANÁLISIS CIENCIOMÉTRICO EN EL DESARROLLO DE TEMAS DE TESIS DE MAESTRÍA EN EL PERIODO 1996-2007. CASO DE ESTUDIO: CENTRO DE INVESTIGACIÓN EN COMPUTACIÓN DEL INSTITUTO POLITECNICO NACIONAL”, al Instituto Politécnico Nacional para su difusión, con fines académicos y de investigación.

Los usuarios de la información no deben reproducir el contenido textual, gráficas o datos del trabajo sin el permiso expreso del autor y/o director del trabajo. Este puede ser obtenido escribiendo a la siguiente dirección [matlacuevat@hotmail.com](mailto:matlacuevat@hotmail.com). Si el permiso se otorga, el usuario deberá dar el agradecimiento correspondiente y citar la fuente del mismo.

*Vazquez Gallo Maribel.*

Maribel Vazquez Gallo

---

## AGRADECIMIENTOS

Es importante comenzar por Dios, a Él le agradezco por haberme dado la fortaleza para cerrar este ciclo y a mi madre por sus oraciones, bendiciones y todo su amor, acompañándome en todo momento.

A la Maestra Martha Durán y a la Maestra Blanca Miranda, por las facilidades otorgadas.

A mis asesores Dra. Pilar Longar Blanco y al Dr. Adolfo Guzmán Arenas por sus aportaciones para la realización de este trabajo.

Al Dr. Adolfo Guzmán Arenas, al Dr. Gilberto Lorenzo Martínez Luna, a las Dra. Pilar Longar Blanco, a la Dra. Hortensia Gómez Viquez, al Dr. Humberto Merrit Tapia por sus conocimientos, opiniones y consejos para mejorar ésta tesis.

A mi pequeña Matlacueyatl, esa personita que ilumina mi vida y hace mas grande mi corazón.

A mi hermano Celso, que sin saberlo he aprendido muchas cosas de él.

Finalmente a la persona que ha pesar de todo siempre me ha apoyado, me alienta día a día a ser mejor persona, gracias amor por la paciencia y la confianza que me tienes para continuar viviendo juntos los sueños y metas que se presentan.

Y a todos los que, de alguna manera, contribuyeron a la elaboración de este trabajo y que no podría listar sin dejar de mencionar alguno.

---

## CONTENIDO

ÍNDICE DE GRÁFICAS .....	VII
ÍNDICE DE FIGURAS .....	VIII
ÍNDICE DE TABLAS .....	X
SIGLAS .....	XI
GLOSARIO.....	XII
RESUMEN .....	1
ABSTRACT .....	2
INTRODUCCIÓN.....	3
<b><u>CAPITULO 1. MARCO TEÓRICO.....</u></b>	<b>9</b>
<b>1.1 BIBLIOMETRÍA.....</b>	<b>9</b>
<b>1.2 CIENCIOMETRÍA .....</b>	<b>10</b>
<b>1.3 MAPAS TECNOLÓGICOS .....</b>	<b>12</b>
<b>1.4 ANÁLISIS DE CITAS.....</b>	<b>13</b>
1.4.1 Mapas basados en Análisis de Citas .....	14
<b>1.5 ANÁLISIS DE PALABRAS ASOCIADAS .....</b>	<b>16</b>
1.5.1 Mapas basados en Análisis de Palabras Asociadas.....	16
<b>1.6 SOFTWARE INFORMÁTICO APLICADO A LA CIENCIOMETRÍA Y BIBLIOMETRÍA .....</b>	<b>17</b>
<b>1.7 APLICACIONES DE CIENCIOMETRÍA Y BIBLIOMETRÍA .....</b>	<b>19</b>
1.7.1 Estudio 1: La bibliometría en la Producción de Autores cubanos en las revistas sobre ciencias de la computación registradas en el Journal Citation Reporter en el periodo 1990-2005” .....	20
1.7.2 Estudio 2: Un análisis para visualizar la historia de la ciencia .....	20
1.7.3 Estudio 3: Como identificar campos de investigación emergentes usando la ciencia: un ejemplo en el campo de la seguridad de la información.....	21
<b><u>CAPITULO 2. APLICACIÓN DE LA BIBLIOMETRÍA .....</u></b>	<b>23</b>
<b>2.1 BASE DE DATOS DE LOS TEMAS DE TESIS.....</b>	<b>23</b>
<b>2.2 PRINCIPALES INDICADORES BIBLIOMÉTRICOS.....</b>	<b>24</b>
2.2.1. Resultados y discusión de Indicadores de Publicación.....	24
<b>2.3 INDICADORES DE PRODUCCIÓN CIENTÍFICA EN TESIS CONCLUIDAS .....</b>	<b>25</b>
2.3.1 Producción de Tesis.....	25
2.3.2 Indicador de Producción.....	26
2.3.3 ley de Lotka .....	31
2.3.4 Ley de Zipf.....	34
<b>2.4 RESULTADOS Y DISCUSIÓN DE INDICADORES DE CITACIÓN .....</b>	<b>38</b>
2.4.1 Apareo Bibliográfico.....	40
2.4.2 Análisis de citas.....	43
2.4.4 Índice de Impacto.....	51
<b><u>CAPITULO 3. APLICACIÓN DEL MÉTODO DE LAS PALABRAS ASOCIADAS .....</u></b>	<b>52</b>
<b>3.1 SELECCIÓN DE METODOLOGÍA DE ANÁLISIS .....</b>	<b>52</b>
3.1.1 Organizando las palabras claves en un ambiente de búsqueda en la web: Una metodología basada en el análisis de palabras asociadas .....	52
3.1.2 Como identificar campos de investigación emergentes usando ciencia: Un ejemplo en el campo de la seguridad de la información. ....	53
<b>3.2 DESCRIPCIÓN DE LA METODOLOGÍA UTILIZADA. ....</b>	<b>54</b>
3.2.1 Decidir el texto y el área a analizar.....	55
3.2.2 Palabras del área a buscar en tesis. ....	56
3.2.3 Registrar en HistCite versión 8.5.26 cada una de las tesis y exportar a Excel.....	56
3.2.4 Normalizar las palabras .....	56
3.2.5 Organizar y almacenar por renglón cada una de las tesis en un archivo.....	57
3.2.6 Dar el archivo con las 90 tesis a Redes 2005 .....	58
3.2.7 Obtener Matrices de adyacencia, normalizadas, mapa estratégico y grafo. ....	62
3.2.8 Mostrar Mapas.....	64

---

3.2.9 Interpretar los Resultados.....	65
<b>3.3 DESCRIPCIÓN DEL MÉTODO DE LAS PALABRAS ASOCIADAS CON ALGORITMO DE AGRUPACIÓN SOBRE CENTROS SIMPLES, UTILIZADO POR REDES 2005.</b> .....	65
3.3.1 Selección de Parámetros para análisis en diez tesis.....	65
3.3.1.1 Interpretación del tema desarrollo en el diagrama estratégico.....	76
<b>3.4 ANÁLISIS DE TÓPICOS DE LAS CIENCIAS DE LA COMPUTACIÓN EN DOCUMENTOS DE TESIS.</b> .....	76
3.4.1 Análisis del Período 1996-2001.....	77
3.4.2 Análisis del Período 2002-2004.....	80
3.4.3 Análisis del Período 2005-2007.....	84
<b><u>CAPITULO 4. BÚSQUEDA DE PATRONES Y TENDENCIAS EN LOS GRAFOS Y DIAGRAMAS ESTRATÉGICOS.</u></b> .....	<b>88</b>
4.1 METODOLOGÍA DE BÚSQUEDA DE PATRONES Y TENDENCIAS .....	88
4.2 ANÁLISIS DE LOS DIAGRAMAS ESTRATÉGICOS Y GRAFOS EN LOS PERIODOS .....	89
<b>CONCLUSIONES Y RECOMENDACIONES.....</b>	<b>92</b>
<b>LÍNEAS DE INVESTIGACIÓN A FUTURO .....</b>	<b>94</b>
<b>BIBLIOGRAFÍA .....</b>	<b>95</b>
<b>ANEXO 1. BASE DE DATOS EN LOS TEMAS DE TESIS.....</b>	<b>98</b>
<b>ANEXO 2. ESTADO Y TRATAMIENTO DE CITAS EN LOS DOCUMENTOS DE TESIS DE MAESTRÍA.....</b>	<b>104</b>
<b>ANEXO 3 CONCEPTOS DE INDICADORES DE CITACIÓN .....</b>	<b>107</b>
<b>ANEXO 4. DESCRIPCIÓN DE LA METODOLOGÍA GENERAL PARA EL ANÁLISIS CIENCIOMÉTRICO .....</b>	<b>113</b>
<b>ANEXO 5. VOCABULARIO EN LOS TÍTULOS DE LAS TESIS.....</b>	<b>116</b>
<b>ANEXO 6. ACM COMPUTING CLASSIFICATION SYSTEM. 1998 VERSION.....</b>	<b>121</b>

---

## ÍNDICE DE GRÁFICAS

GRÁFICA 2.3.1. 1 PRODUCCIÓN DE TESIS SOBRE CIENCIAS DE LA COMPUTACIÓN EN EL PERIODO 1996-2007 ..	26
GRÁFICA 2.3.2. 1 PRODUCCIÓN DE TESIS POR LABORATORIO .....	27
GRÁFICA 2.3.2. 2 PRODUCTIVIDAD POR DIRECTOR DE TESIS.....	28
GRÁFICA 2.3.2. 3 PORCENTAJE DE TESIS CON DOS DIRECTORES .....	28
GRÁFICA 2.3.2. 4 CODIRECCIONES DE TESIS POR AÑO .....	29
GRÁFICA 2.3.2. 5 RED DE COLABORACIÓN ENTRE LABORATORIOS .....	30
GRÁFICA 2.3.2. 6 PROCEDENCIA DE INVESTIGADORES EXTERNOS .....	31
GRÁFICA 2.3.3. 1 DISTRIBUCIÓN DE LOTKA .....	33
GRÁFICA 2.3.4. 1 LEY DE ZIPF.....	37
GRÁFICA 2.3.4. 2 ZONA DE PALABRAS CON MAYOR CONTENIDO SEMÁNTICO.....	38

---

## ÍNDICE DE FIGURAS

FIGURA 1 CRECIMIENTO DE LAS BASES DE DATOS DE CYT Y SU USO EN LA CIENCIOMETRÍA.....	5
FIGURA 1.2.2.1 LA ROSA DE LOS VIENTOS DE LA INVESTIGACIÓN .....	10
FIGURA 1.4. 1 EJEMPLO. MAPA DE ANÁLISIS DE CITAS .....	15
FIGURA 1.5. 1 DIAGRAMA ESTRATÉGICO. PRINCIPALES TEMAS DE INVESTIGACIÓN.....	17
FIGURA 1.7.2. 1 ARTÍCULOS QUE CITAN A PRICE VERSUS ARTÍCULOS QUE USAN EL TERMINO “CIENCIOMETRÍA”, EN EL PERIODO 1956-2006. ....	21
FIGURA 1.7.3. 1 MAPA GENERADO A TRAVÉS DEL ANÁLISIS DE PALABRAS CON APARIENCIA SIMULTANEA 1994- 1999 .....	22
FIGURA 2.3.4. 1 COLECCIÓN DE TESIS EN HISTCITE.....	34
FIGURA 2.4. 8 RECURSOS DE INFORMACIÓN.....	46
FIGURA 2.4. 9 PALABRAS CLAVE EN TÍTULOS DE LIBROS Y TESIS CITADOS .....	47
FIGURA 2.4. 10 IDIOMA DE LAS CITACIONES .....	48
FIGURA 2.4. 11 POR PAÍS .....	48
FIGURA 2.4. 12 PUBLICACIÓN POR AÑO DE TESIS Y LIBROS CITADOS .....	48
FIGURA 2.4. 13 AUTOR <i>GOLDFARB</i> CON TRES CITAS .....	49
FIGURA 2.4. 14 TRES TEMAS DE TESIS QUE CITAN A <i>GOLDFARB</i> .....	50
FIGURA 2.4. 15 CO-CITACIÓN EN DOS AUTORES POR PARTE DE TRES TEMAS DE TESIS .....	50
FIGURA 2.4. 16 AUTOR <i>BOOCH G</i> CON DOCE PUBLICACIONES Y QUINCE CITAS .....	51
FIGURA 3.1. 1 METODOLOGÍA DE TRABAJO PARA FORMAR TESAURUS EN FORMA AUTOMÁTICA .....	53
FIGURA 3.1. 2 METODOLOGÍA DE TRABAJO PARA LOCALIZAR TENDENCIAS EN UN CAMPO DE CONOCIMIENTO.53	
FIGURA 3.2. 1 METODOLOGÍA DE TRABAJO PARA LOCALIZAR PATRONES Y TENDENCIAS EN TESIS DE MAESTRÍA .....	55
FIGURA 3.2.5. 1 FORMATO PARA ALMACENAR EN REDES 2005 .....	57
FIGURA 3.2.6. 1 SELECCIÓN DEL PROGRAMA REDES 2005.....	58
FIGURA 3.2.6. 2 SELECCIÓN DE NUEVO PROYECTO. ....	58
FIGURA 3.2.6. 3 ELEGIR NOMBRE DEL PROYECTO.....	59
FIGURA 3.2.6. 4 ELECCIÓN DE ORIGEN DE DATOS O ARCHIVO DE DATOS. ....	59
FIGURA 3.2.6. 5 ASIGNACIÓN DE PARÁMETROS.....	60
FIGURA 3.2.6. 6 NÚMERO DE PALABRAS EN ANÁLISIS Y PALABRAS QUE SUPERAN EL UMBRAL. ....	60
FIGURA 3.2.6. 7 LISTADO DE LAS PALABRAS QUE SE LOCALIZAN EN EL ANÁLISIS. ....	61
FIGURA 3.2.7. 1 MATRIZ DE ADYACENCIA. CO-OCURRENCIA DE PALABRAS.....	62
FIGURA 3.2.7. 2 MATRIZ NORMALIZADA CON EL ÍNDICE DE EQUIVALENCIA. ....	62
FIGURA 3.2.7. 3 LISTA DE ENLACES ORDENADOS POR ÍNDICE DE EQUIVALENCIA. ....	63
FIGURA 3.2.7. 4 FICHA INFORMATIVA DE UNO DE LOS AGRUPAMIENTOS FORMADOS .....	64
FIGURA 3.2.8. 1 DIAGRAMA ESTRATÉGICO Y GRAFO DE UNO DE LOS AGRUPAMIENTOS DEL PERIODO 1996-2007. .....	64
FIGURA 3.3. 1 PARÁMETROS DEL ANÁLISIS .....	66
FIGURA 3.3. 2 NÚMERO DE PALABRAS QUE SUPERAN EL UMBRAL .....	66
FIGURA 3.3. 3 LISTADO DE PALABRAS QUE SUPERAN EL UMBRAL .....	67

---

FIGURA 3.3. 4 MATRIZ DE ADYACENCIA O CO-OCURRENCIA CON PALABRAS. ....	67
FIGURA 3.3. 5 MATRIZ NORMALIZADA CON ÍNDICE DE EQUIVALENCIA, PERIODO 1996-2001 .....	68
FIGURA 3.3. 6 LISTA DE ENLACES ORDENADOS POR ÍNDICE DE EQUIVALENCIA .....	69
FIGURA 3.3. 7 MATRIZ CON RESULTADO PARCIAL DEL ALGORITMO DE AGRUPACIÓN POR CENTROS SIMPLES ..	69
FIGURA 3.3. 8 PALABRA CENTRAL “DESARROLLO” .....	70
FIGURA 3.3. 9 MATRIZ CON RESULTADO PARCIAL DEL ALGORITMO DE AGRUPACIÓN POR CENTROS SIMPLES. .	70
FIGURA 3.3. 10 FIGURA EN FORMA DE ESTRELLA. ....	71
FIGURA 3.3. 11 SEÑALAMIENTO DE LAS PALABRAS DE LOS ENLACES DEL DESCRIPTOR “DISEÑO” .....	72
FIGURA 3.3. 12 VISUALIZACIÓN DE ENLACES DEL DESCRIPTOR “DISEÑO” .....	72
FIGURA 3.3. 13 VISUALIZACIÓN DE ENLACES DEL DESCRIPTOR “USUARIO”. ....	73
FIGURA 3.3. 14 DIAGRAMA ESTRATÉGICO Y GRAFO DEL AGRUPAMIENTO DEL PERIODO 1996-2001.....	73
FIGURA 3.3. 15 FICHA INFORMATIVA DEL AGRUPAMIENTO DEL TEMA <i>DESARROLLO</i> .....	74
FIGURA 3.4.1. 1 DIAGRAMA ESTRATÉGICO DEL PERIODO 1996-2001 .....	77
FIGURA 3.4.1. 2 GRAFO DEL TEMA “ <i>SISTEMA</i> ” DEL PERIODO 1996-2001.....	78
FIGURA 3.4.1. 3 GRAFO DEL TEMA “ <i>PROGRAMA</i> ” DEL PERIODO 1996-2001 .....	79
FIGURA 3.4.1. 4 GRAFO DEL TEMA “ <i>SERVIDOR</i> ” DEL PERIODO 1996-2001 .....	79
FIGURA 3.4.1. 5 GRAFO DEL TEMA “ <i>DESARROLLO</i> ” DEL PERIODO 1996-2001 .....	79
FIGURA 3.4.1. 6 GRAFO DEL TEMA “ <i>PROCESO</i> ” DEL PERIODO 1996-2001 .....	80
FIGURA 3.4.2. 1 DIAGRAMA ESTRATÉGICO DEL PERIODO 2002-2004 .....	81
FIGURA 3.4.2. 2 GRAFO DEL TEMA “ <i>TECNOLOGÍA</i> ” DEL PERIODO 2002-2004 .....	82
FIGURA 3.4.2. 3 GRAFO DEL TEMA “ <i>MODELO</i> ” DEL PERIODO 2002-2004.....	82
FIGURA 3.4.2. 4 GRAFO DEL TEMA “ <i>SISTEMA</i> ” DEL PERIODO 2002-2004.....	83
FIGURA 3.4.2. 5 GRAFO DEL TEMA “ <i>SERVIDOR</i> ” DEL PERIODO 2002-2004.....	83
FIGURA 3.4.2. 6 GRAFO DEL TEMA <i>USUARIO</i> DEL PERIODO 2002-2004.....	83
FIGURA 3.4.3. 1 DIAGRAMA ESTRATÉGICO DEL PERIODO 2005-2007 .....	84
FIGURA 3.4.3. 2 GRAFO DEL TEMA “ <i>USUARIO</i> ” DEL PERIODO 2005-2007.....	85
FIGURA 3.4.3. 3 GRAFO DEL TEMA “ <i>SISTEMA</i> ” DEL PERIODO 2005-2007.....	86
FIGURA 3.4.3. 4 GRAFO DEL TEMA “ <i>COMUNICACIÓN</i> ” DEL PERIODO 2005-2007 .....	86
FIGURA 3.4.3. 5 GRAFO DEL TEMA “ <i>PROCESO</i> ” DEL PERIODO 2005-2007.....	86
FIGURA 3.4.3. 6 GRAFO DEL TEMA “ <i>ALGORITMO</i> ” DEL PERIODO 2005-2007 .....	87
FIGURA 4.2. 1 <i>USUARIO</i> COMO DESCRIPTOR DEL TEMA <i>PROGRAMA</i> . ....	90
FIGURA 4.2. 2 <i>USUARIO</i> COMO TEMA EN EL CUARTO CUADRANTE. ....	90
FIGURA 4.2. 3 <i>USUARIO</i> COMO TEMA EN EL PRIMER CUADRANTE.....	91

---

## ÍNDICE DE TABLAS

TABLA 1 VOLÚMENES DE REGISTROS EN BASES DE DATOS ESPECIALIZADAS.....	3
TABLA 1.2.2. 1 INDICADORES BIBLIOMÉTRICOS UTILIZADOS EN CIENCIOMETRÍA.....	11
TABLA 1.6 1 PRINCIPALES HERRAMIENTAS ESTADÍSTICAS UTILIZADAS EN CIENCIOMETRÍA.....	18
TABLA 2.2. 1 INDICADORES DE PUBLICACIÓN.....	24
TABLA 2.2. 2 INDICADORES DE PRODUCCIÓN EN TESIS CONCLUIDAS.....	25
TABLA 2.3.2. 1 MATRIZ DE COLABORACIÓN EN DIRECCIÓN DE TESIS ENTRE LABORATORIOS.....	29
TABLA 2.3.3. 1 DISTRIBUCIÓN DE FRECUENCIAS POR DIRECCIONES DE TESIS.....	32
TABLA 2.3.3. 2 DISTRIBUCIÓN DE FRECUENCIAS POR DIRECCIONES DE TESIS.....	33
TABLA 2.3.3. 1 DISTRIBUCIÓN DE FRECUENCIAS POR DIRECCIONES DE TESIS.....	32
TABLA 2.3.3. 2 DISTRIBUCIÓN DE FRECUENCIAS POR DIRECCIONES DE TESIS.....	33
TABLA 2.3.4. 1 VOCABULARIO EN LOS TÍTULOS DE LAS TESIS.....	35
TABLA 2.3.4. 2 PALABRAS CON MAYOR FRECUENCIA EN LOS TÍTULOS DE LAS TESIS.....	36
TABLA 2.3.4. 3 VOCABULARIO ORDENADO POR FRECUENCIAS.....	37
TABLA 2.4. 1 INDICADORES DE CITACIÓN.....	39
TABLA 2.4. 2 ANÁLISIS PARA APLICAR LOS INDICADORES DE CITACIÓN.....	40
TABLA 2.4. 3 FRECUENCIAS DE CITAS.....	43
TABLA 3.4.1. 1 PARÁMETROS Y RESULTADOS DE LA EJECUCIÓN CON REDES 2005 EN EL PERIODO 1996-2001	77
TABLA 3.4.2. 1 PARÁMETROS Y RESULTADOS DE LA EJECUCIÓN CON REDES 2005 EN EL PERIODO 2002-2004	80
TABLA 3.4.3. 1 PARÁMETROS Y RESULTADOS DE LA EJECUCIÓN CON REDES 2005 EN EL PERIODO 2005-2007. .....	84
TABLA 4.2. 1 UBICACIÓN DE UN TEMA EN DIAGRAMAS ESTRATÉGICOS.....	89

---

## SIGLAS

<b>SIGLA</b>	<b>SIGNIFICADO</b>
<b>ACM</b>	Association for Computing Machinery
<b>CIC</b>	Centro de Investigación en Computación
<b>CINVESTAV-Z</b>	Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional. Unidad Zacatenco
<b>CONACYT</b>	Consejo Nacional de Ciencia y Tecnología
<b>ESCOM</b>	Escuela Superior en Computación
<b>ESIME-Z</b>	Escuela Superior de Ingeniería Mecánica y Eléctrica. Unidad Zacatenco
<b>FORMATO APA</b>	American Psychological Association (Asociación Psicológica Americana)
<b>IEEE</b>	Institute of Electrical and Electronics Engineers (Instituto de Ingenieros en Electricidad y Electrónica)
<b>IMP</b>	Instituto Mexicano del Petróleo
<b>INAOE</b>	Instituto Nacional de Astrofísica, Óptica y Electrónica
<b>ISI</b>	Institute for Scientific Information (Instituto de Información Científica)
<b>JCR</b>	Journal Citation Report (Reporte de Citas a Revistas)
<b>SIP</b>	Secretaria de Investigación y Posgrado
<b>SCI</b>	Science Citation Index (Índice de Citas de la Ciencia)
<b>SCCI</b>	Social Science Citation Index (Indice de Citas en Ciencias Sociales)
<b>UNAM</b>	Universidad Nacional Autónoma de México

## GLOSARIO

TERMINO	DESCRIPCIÓN
<b>BASES DE DATOS</b>	Todo tipo de información organizada en datos almacenados en la memoria de una computadora, diseñada para facilitar el acceso de manera estandarizada.
<b>BIBLIOMETRÍA</b>	Es la aplicación de estadísticas en una área de investigación con el objetivo de estudiar y describir la actividad científica
<b>CIENCIOMETRÍA</b>	Parte de la base de que los resultados de las investigaciones científicas y técnicas se plasman en forma escrita, lo que permitirá ser analizada para descubrir su evolución y sus características. (Escorsa y Maspons, 2001).
<b>CITA</b>	Palabras dichas o escritas por otra persona con la finalidad de justificar lo que dice o escribe
<b>DIAGRAMA ESTRATÉGICO</b>	Es un diagrama cartesiano que se divide en cuatro cuadrantes, en el eje X se encuentra la centralidad y en el eje Y la densidad. Sirven para representar, conceptos, autores, laboratorios, entre otros.
<b>EVOLUCIÓN</b>	Desarrollo o transformación de ideas o teorías
<b>GRAFO</b>	Conjunto de vértices V y un conjunto de arcos o aristas A. Los vértices se denominan también nodos o puntos.
<b>HERRAMIENTA</b>	Instrumento para facilitar una tarea mecánica.
<b>INDICADOR</b>	Parámetro que permite evaluar el comportamiento o tendencias de algún tipo de información específica.
<b>INDICADOR BIBLIOMÉTRICO</b>	Datos estadísticos que se deducen de las publicaciones científicas.
<b>INFORMÁTICA</b>	Disciplina encargada del estudio de métodos, procesos, técnicas, desarrollos y su utilización en ordenadores (computadores), con el fin de almacenar, procesar y transmitir información y datos en formato digital.
<b>MAPA HISTORIOGRÁFICO</b>	Visualización grafica construida a partir de citas a través del tiempo
<b>METÓDO DE PALABRAS ASOCIADAS</b>	Detección de palabras que caracterizan un tema y en contar la co-aparición de estas, en campos tales como títulos de artículos, palabras claves de artículos, abstracts o bien directamente del texto libre. (Escorsa y Maspons, 2001)
<b>RECURSO DE INFORMACIÓN</b>	Son documentos escritos por científicos y tecnólogos
<b>REFERENCIA BIBLIOGRAFICA</b>	Conjunto de fichas técnicas que permiten identificar una publicación
<b>SOFTWARE</b>	Conjunto de programas de una computadora u ordenador, que le permiten realizar las funciones asignadas por el usuario.

---

## RESUMEN

En este trabajo se presenta un estudio cuantitativo, donde se utilizan indicadores bibliométricos tales como indicadores de publicación e indicadores de análisis de citas, se hace uso del método de las palabras asociadas (análisis *cowords*) que utiliza la cuantimetría, todo esto en una muestra de documentos de tesis de maestría concluidas en ciencias de la computación del Centro de Investigación en Computación en el periodo de 1996 hasta 2007. Se ha organizado la información necesaria en una base de datos para este análisis cuantitativo, utilizando los datos de hasta 202 tesis; datos que se han recuperado acudiendo a diferentes fuentes de información, como lo son: Biblioteca, Área de Control Escolar y Apoyo de Investigadores.

En la parte de indicadores de publicación se describe la producción científica a partir de los documentos de tesis; en donde se analiza la productividad de los directores en diversos tópicos de las Ciencias de la Computación, además de analizar si se tiene la preferencia o la tendencia en determinados tópicos. Los indicadores de análisis de citas describen como es la fuente de conocimiento en cuanto a citas a libros.

Con el método de las palabras asociadas se generan diagramas estratégicos y grafos (representación visual de la información) permitiendo seguir la evolución a través del tiempo (patrones y tendencias) con los tópicos de las ciencias de la computación para analizar las relaciones entre ellos, los tópicos son de acuerdo al grupo de *IEEE Computing Society and ACM* y profesores del CIC.

Esta tesis deja las bases para realizar comparativos con el conocimiento que reflejan los diagramas estratégicos en el campo de las ciencias de la computación en los diferentes centros de investigación del I.P.N, con base en lo que se escribe en los documentos de tesis. Estos diagramas se pueden enriquecer si se adiciona el análisis de artículos u otra fuente de información.

---

## ABSTRACT

This these presents a study scientometric, where uses bibliometric indicators such as indicators of publication and citation analysis indicators, makes use of the method of associated words (analysis cowords) that use scientometrics, everything on a sample of master's thesis in computer science finished Computer Research Center from 1996 to 2007. It has organized the necessary information in a database for the scientometric analysis, using data up to 202 theses, data to be retrieved by going to different sources of information, such as: Library, School Area Control and Support of Investigators.

The publication indicators described the scientific production in the thesis documents; where analyze the productivity of managers in various topics of Computer Science, in addition to analyzing whether there was any preference or tendency in certain topics. Indicators of citation analysis described how the source of knowledge regarding references is.

With the method of associated words are generated strategic diagrams and graphs (visual representation of information) allowing to follow the evolution over time (patterns and trends) with the topics of computer science to analyze the relationships between them, topics are according to the group of IEEE Computing Society and ACM and teachers of CIC.

This thesis leaves the basis for do the comparison with the knowledge that reflect the strategic diagrams in this field of computer science at the different research centers at the I.P.N and what actually is studied in them. These diagrams could be enrich if add the analysis of articles or other information source.

## INTRODUCCIÓN

Hoy en día se generan elevados volúmenes de información derivados de la actividad científica; como necesidad se impone analizarla para poder obtener conclusiones sobre su avance, evolución, generación de un nuevo conocimiento, si es que existe, análisis que es utilizado por las organizaciones tanto de gobierno como empresas privadas quienes buscan decidir sobre gestión de políticas en ciencia y tecnología, lo que implica analizar grandes volúmenes de información que se encuentran en bases de datos especializadas Deleus (2003), Van (2003), Zhu (1999), Porter (1999). Algunos ejemplos de estas bases de datos con sus volúmenes. Tabla 1.

El analizar los grandes y complejos volúmenes de información de la actividad científica se ha llevado a cabo de una forma sistematizada y formal con la ayuda de la bibliometría, cienciometría y la informetría. De estas tres disciplinas, la informetría aún se encuentra en estado de definición, por lo cual en este estudio no se tomará en consideración.

**Tabla 1 Volúmenes de registros en bases de datos especializadas**

No.	Nombre de Base de datos	Número de revistas	Tipo de revista	Producción referencias anuales	Número de registros
<b>BASES DE DATOS INTERNACIONALES</b>					
1	SCI (Science Citation Index)	En cd rom 3500 y on line 5300	Ciencias de la vida, medio ambiente, tecnología, física, química y medicina.	900.000	-----
2	MEDLINE (Index Medicus)	3200	Revistas sobre medicina	400.000	8 millones
3	PASCAL (Bibliographie Internacioanl)	6500	Revistas de medicina, medio ambiente, ciencias de la vida y tecnología	-----	14.7 millones
4	BIOSIS (Biological Abstract)	9.000	Revistas del campo de la ciencia de la vida		600.000
5	CA (Chemical Abstract)	10.000	Revistas en el área de química, bioquímica, e ingeniería química incluyendo disciplinas básicas y aplicadas	700.000	-----
6	CAB Internacional	11.000	Revistas de nutrición, bosques y suelo, agricultura y veterinaria	-----	-----

7	COMPENDEX (Engineering Index)	5.000	Revistas de ingeniería e informática	200.000	-----
8	INSPEC (Physics Abstracts)	4.000	Revistas de astronomía, física e ingeniería eléctrica, tecnologías de la información e informática	300.000	-----
9	ICYT (Indice Español de Ciencia y Tecnología)	550	Revistas españolas de ciencias de la vida, tecnología, ciencias exactas y naturales, agronomía y ciencias de la tierra y el espacio	6500	10,000,000
10	MathSci (Mathematical Science)	3.100	Revistas de matemáticas, estadística, computación y sus aplicaciones en mecánica, física, ingeniería, medicina, biología, agricultura, educación, etc.	60.000	-----
11	SCOPUS		Revistas de Agricultura, Biología, Química, Geología, Economía, Negocios, Ingeniería, Salud, Ciencias de la vida, Matemáticas, Física, Psicología y Ciencias Sociales.	-----	17 millones
<b>BASES DE DATOS REGIONALES</b>					
11	LILIACS (Literatura de América Latina y el Caribe en Ciencias de la Salud)	670	Revistas de literatura científica y tecnológica en el campo de las ciencias medicas.	-----	150
12	CLASE	1200	Revistas latinoamericanas especializadas en ciencias sociales y humanidades	-----	160.000
13	PERIODICA	1300	Revistas latinoamericanas en ciencia y tecnología	-----	160.000
14	AGRIS	-----	Revistas en el área de Agricultura y medio ambiente	-----	1.900.000

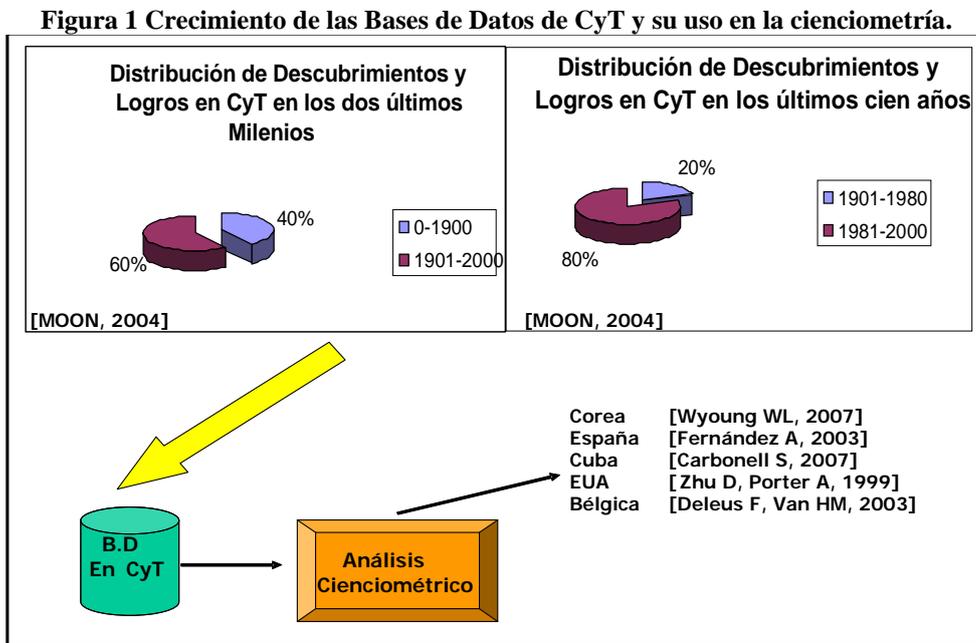
Fuente: De Filippo y Fernández. (n.d)

Moon, (2004) citando a Woo, (2007) dio una explicación del porque de los crecientes volúmenes de información sobre logros y descubrimientos científicos, indicando que del 100% de los logros y descubrimientos alrededor del 60% se han llevado a cabo en los

últimos 100 años y del 60% de estos logros y descubrimientos el 80% se han llevado a cabo en los últimos 20 años.

Varios países en el mundo como Cuba, Corea y España realizan la evaluación de la actividad científica, siendo un elemento imprescindible para que los programas de investigación, tecnología y desarrollo se implementen en una sociedad. Figura 1.

La evaluación ha sido formalizada con la disciplina llamada *cienciometría*, por medio de un proceso sobre las bases de datos especializadas y ha contribuido al desarrollo de indicadores que constituyen herramientas clave en la gestión de la política científica y tecnológica, y en los procesos de toma de decisiones (Larreina, Hernando y Grisaleña: 2006).



Fuente: Elaboración propia a partir de Moon.2004

Por lo anterior en un centro de investigación puede ser de gran utilidad hacer una revisión de su producción científica por medio de la *cienciometría*, con el objetivo de desarrollar una planeación para acelerar y mejorar el crecimiento de la producción científica con la finalidad de preveer o mejorar las estrategias que lleven a cabo los objetivos de la organización (Fernández et al., 2003).

Entre los objetivos del Centro de Investigación en Computación es divulgar los resultados de su producción científica por medio de documentación escrita, en este trabajo, el interés se centra sobre los documentos de las tesis y que se pueden ver como un producto final del

Centro, con los cuales es posible evaluar su actividad científica en sus áreas de investigación.

En México, en muchos centros de investigación no se tienen elementos organizados de tal manera que permitan analizar su producción científica que sea más allá de realizar estadísticas sobre esta. Un análisis más profundo, es el poder obtener descripciones más precisas de su producción, ya sea clasificada por laboratorios, porcentaje de tesis codirigidas, colaboración con investigadores externos, principales áreas de trabajo, áreas más desarrolladas, el poder realizar pronósticos sobre el desarrollo de sus áreas, entre otras descripciones.

Los documentos de tesis representan una “materia prima” que hay que procesar para obtener información útil y que puede ayudar a resolver las descripciones del párrafo anterior, este proceso es llevar a cabo una investigación en el campo de la ciencia y la tecnología.

El mantenimiento de una producción científica pertinente de un centro de investigación con desarrollo tecnológico se debe mantener para cumplir los requisitos mínimos requeridos por las entidades del gobierno responsables de otorgar recursos como CONACYT (Consejo Nacional de Ciencia y Tecnología) y la SIP (Secretaría de Investigación y Posgrado), y en esa medida sigan financiando la actividad científica, la cual requiere una evaluación mediante el análisis cuantitativo formal. El análisis formal no es una tarea sencilla, pues se requiere desde una simple estadística hasta modelos analíticos complejos (Salvador, 2005).

Por la bibliografía que se ha analizado hasta el momento no se encuentra evidencia de que se exista una revisión cuantitativa en la disciplina de las Ciencias de la Computación en México, por ello mismo se motiva a llevar a cabo este proyecto de investigación.

En virtud de lo anterior, el presente trabajo se orienta al análisis cuantitativo en los documentos de tesis en la disciplina o campo de las Ciencias de la Computación; con lo cual el estudio busco responder algunas de las siguientes interrogantes:

- ¿Cuáles son los tópicos de las ciencias de la computación que se estudian realmente en los temas de tesis?
- ¿En qué tópicos de las ciencias de la computación los temas de tesis se relacionan?
- Estas relaciones entre los tópicos de los temas de tesis ¿Generan nuevos tópicos?
- ¿En que tópicos se están generando especialistas?

El objetivo de este trabajo es utilizar la cienciometría para identificar la evolución de los tópicos en la disciplina de las ciencias de la computación, basándose en la información del desarrollo de los temas de tesis del Centro de Investigación en Computación del I.P.N, el cual se cumplió estudiando la producción de tesis de maestría realizadas, aportando una visión diacrónica de las mismas, realizando un análisis de las fuentes de conocimiento de los temas de tesis del período en referencia y aplicar el método de las palabras asociadas en los títulos y resúmenes de los documentos de tesis para localizar patrones y tendencias en los tópicos de las ciencias de la computación.

Una de las aportaciones a este trabajo de investigación es proporcionar una revisión detallada de la aplicación metodológica de la cienciometría, describiendo desde la selección de objetivo, alcance, mostrando los cálculos necesarios y la interpretación de los diagramas estratégicos, junto con los grafos correspondientes. La aplicación detallada en la disciplina de las ciencias de la computación en un centro de investigación en México, pero se puede aplicar de manera general a cualquier otro centro con características similares.

Este trabajo esta dividido en cuatro capítulos, el primer capítulo corresponde al marco teórico en dónde se dan conceptos y definiciones de algunos autores pioneros en bibliometría, cienciometría y análisis de palabras asociadas, el software informático que es aplicado en estudios cienciométricos, destacando casos de estudio de estas disciplinas.

En el capítulo dos corresponde a la aplicación de la bibliometría, que describen los indicadores de producción e indicadores de análisis de citas en las tesis de maestría. En el capítulo tres corresponde a la aplicación del método de palabras asociadas y el último a la búsqueda de patrones y tendencias en los grafos y diagramas estratégicos. La última parte corresponde a las conclusiones y líneas de investigación. También se adicionan cinco anexos, el primero muestra la información de la base de datos, en el segundo se explica el porque se decidió solo realizar el análisis de citas a libros en las tesis concluidas de

maestría, el tercero describe la información en detalle de los indicadores de publicación y de análisis de citas, el cuarto corresponde a la descripción de la metodología para el análisis cuantitativo y el último muestra el vocabulario en los títulos de las tesis.

# Capítulo 1.

## Marco Teórico

En este capítulo se abordarán los fundamentos teóricos principales de bibliometría y Cienciometría.

### 1.1 Bibliometría

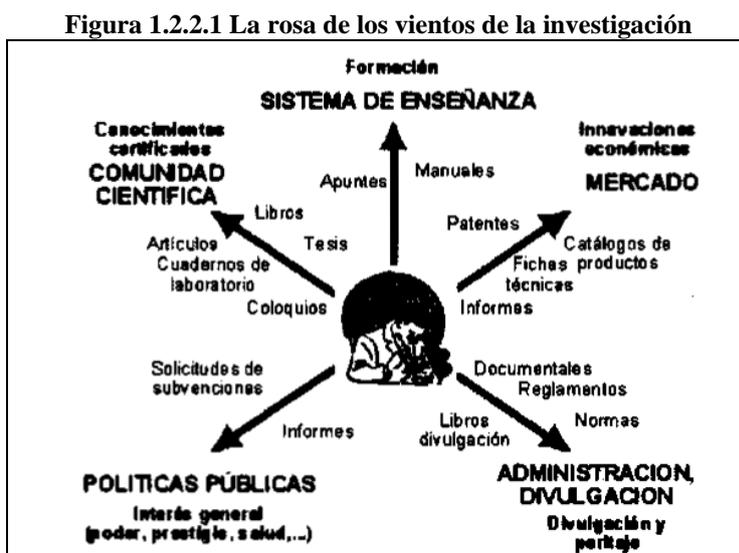
De acuerdo con Callon (1993), la bibliometría se interesa principalmente por los problemas de gestión de las bibliotecas y los centros de documentación, lo que incluye el recuento de artículos y publicaciones así como estudios más detallados sobre una disciplina determinada.

Algunas definiciones de bibliometría son las siguientes:

- a) La bibliometría es la ciencia que estudia la naturaleza y el curso de una disciplina (en cuanto da lugar a publicaciones) mediante la estadística y el análisis de las diversas facetas de la comunicación escrita (Escorsa y Maspons, 2001 citando a Pritchard, 1969).
- b) La bibliometría es la explotación estadística de las publicaciones (Escorsa y Maspons, 2001 citando a Quoniam, 1992).
- c) La bibliometría es una herramienta basada en la aplicación de métodos estadísticos y matemáticos que tiene por objeto facilitar la comparación y comprensión de conjuntos de referencias bibliográficas (Escorsa y Maspons, 2001 citando a Rostaing, 1993).

## 1.2 Cienciometría

La cienciometría parte de la base de que los resultados de las investigaciones científicas y técnicas se plasman en forma escrita a través de cinco dimensiones principales de acuerdo al Centro de Sociología de la Innovación de la Escuela de Minas de París denominada la rosa de los vientos de la investigación. Figura 1.2.2.1 Así pues, el progreso científico y técnico queda registrado en escritos, lo que permitirá ser analizada para descubrir su evolución y sus características (Escorsa y Maspons, 2001)



Fuente: Escorsa y Maspons. 2001

Algunas definiciones de la cienciometría son las siguientes (Escorsa y Maspons, 2001):

- La cienciometría designa aquellos trabajos dedicados al análisis cuantitativo de la actividad científica y técnica (Callon y otros, 1993). Se dedica exclusivamente al análisis de los documentos redactados por los investigadores y los técnicos (artículos, patentes). De alguna manera, intenta identificar las leyes que rigen la actividad científica.
- La cienciometría comprende el conjunto de trabajos dedicados al análisis cuantitativo de la investigación científica y técnica (Callon y otros, 1993).
- La cienciometría incluye la bibliometría, como una herramienta entre otras, para analizar los sistemas de investigación (Rostaing, 1993)
- Es "el estudio de la medición del progreso científico y tecnológico" (Garfield, 1979b).

Se hace una separación entre ambas disciplinas, la bibliometría se basa en el estudio cuantificables de libros, documentos, revistas, artículos, autores; obteniendo con esto estadísticas, mientras que la cienciometría identifica disciplinas de interés utilizando métodos formales, basándose en el análisis cuantitativo de determinados indicadores bibliométricos, como son autores de artículos, citas que sobresalen en la bibliografía de cada artículo, palabras contenidas en los títulos de los artículos o en los resúmenes, estos indicadores se encuentran presentes en las bases de datos, en la Tabla 1.2.2.1 se muestran los principales indicadores bibliométricos (Escorsa y Maspons, 2001):

**Tabla 1.2.2. 1 Indicadores bibliométricos utilizados en cienciometría**

<b>Indicadores de Actividad Científica</b>	
<ul style="list-style-type: none"> <li>• Número y distribución de publicaciones (artículos, patentes, libros) a escala internacional.</li> <li>• Publicaciones producidas por centros de investigación y países.</li> <li>• Productividad de los autores (ley de lotka).</li> <li>• Opiniones de expertos.</li> </ul>	
<b>Indicadores de Impacto</b>	
<ul style="list-style-type: none"> <li>• Número de citas recibidas.</li> <li>• Medición del factor de impacto de los trabajos de investigación</li> </ul>	
<b>Indicadores Relacionales</b>	
Primera Generación	Segunda Generación
<ul style="list-style-type: none"> <li>• Análisis de las citas comunes (cocitaciones).</li> <li>• Estudio de patentes:               <ul style="list-style-type: none"> <li>○ Citas de patentes en artículos científicos.</li> <li>○ Publicaciones de artículos científicos por inventores</li> <li>○ Registro de patentes por investigadores</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• Análisis de palabras conjuntas.</li> <li>• Estudio de patentes a partir de la coocurrencia de palabras.               <ul style="list-style-type: none"> <li>○ Áreas de investigación.</li> <li>○ Organizaciones involucradas.</li> <li>○ Colaboración entre organizaciones y países.</li> </ul> </li> </ul>

**Fuente: Escorsa y Maspons. 2001**

Por medio de estos indicadores se puede determinar (Escorsa y Maspons, 2001):

- a) El crecimiento de cualquier campo de la ciencia, según la variación cronológica del número de trabajos publicados en él.
- b) El nivel de envejecimiento de los campos científicos, en función de la “vida media” de las referencias de sus publicaciones.
- c) La evolución cronológica de la producción científica en función del año de la publicación de los documentos.
- d) La productividad de los autores, medida por el número de sus trabajos publicados.
- e) La colaboración entre los científicos o instituciones, medida por el número de autores por trabajo o centros de investigación que colaboran.
- f) El impacto o visibilidad de las publicaciones dentro de la comunidad científica internacional, medido por el número de citas que reciben éstas por parte de trabajos posteriores.

- g) El análisis y evaluación de las fuentes difusoras de los trabajos, por medio de indicadores de impacto de las fuentes. La publicación de un artículo en una revista de prestigio recibe más citas que si se publicará en una revista con un prestigio menor.
- h) Las trayectorias tecnológicas seguidas por empresas o países en un periodo determinado de tiempo, de acuerdo con su actividad patentadora.

La Tabla 1.2.2.1 muestra los indicadores de actividad científica y de impacto, los cuáles proporcionan información específica sobre el volumen y el impacto de las actividades de investigación científicas, mientras que los indicadores relacionales persiguen conocer las relaciones y las interacciones entre los diferentes elementos bibliográficos (investigadores, sectores, campos tecnológicos, etc.). Estos indicadores relacionales se clasifican en cocitaciones o citas conjuntas (de primera generación) y análisis de palabras conjuntas (de segunda generación). (Escorsa y Maspons, 2001)

Una forma de representar la información científica de una manera diferente a la tradicional, visualizando los datos permitiendo hacer una abstracción de estos y ver aquella información que se encuentra oculta, es por medio de los mapas tecnológicos, que se explicara en el siguiente apartado.

### **1.3 Mapas Tecnológicos**

Los mapas tecnológicos son representaciones gráficas que permiten visualizar los avances tecnológicos que están teniendo lugar, además de proporcionar una idea de cómo evolucionará una determinada tecnología a lo largo del tiempo. Estos mapas permiten detectar tecnologías emergentes y por lo consiguiente nuevas oportunidades de negocio para la empresa, sin embargo para apreciar su evolución resulta muy útil el comparar un mapa con mapas anteriores (Escorsa y Valls, 1997).

Los pasos para la elaboración de un mapa tecnológico son los siguientes (Escorsa, Maspons, Rodríguez, 2001):

- a) Determinación de la estrategia de búsqueda conjuntamente con un experto en la materia (que asesora sobre revistas que cubren el área, palabras clave, etc).
- b) Selección de fuentes de información que cubren mejor el área objeto de análisis.

- c) Teledescarga de la información en la computadora. Consiste en introducción de los registros de la o las bases de datos *on line*, en CD-ROM seleccionadas, en la memoria de la computadora que realizará el tratamiento de los datos.
- d) Homogeneización de la información, que tenga en cuenta las especificidades de cada base y cada formato.
- e) Construcción del *descriptor* de la base de datos, que retiene únicamente los campos que se utilizarán posteriormente.
- f) Depuración de la información (confección de un diccionario de términos sinónimos, eliminación de registros duplicados).
- g) Obtención de los listados y de las *matrices de coocurrencia* deseadas (palabras-palabras, empresas-códigos de clasificación de patentes, palabras clave-empresas, etc.).
- h) Obtención de los mapas tecnológicos.
- i) Validación de los resultados con la ayuda de un experto en la materia.

Para la elaboración de los mapas tecnológicos se hace uso de un nuevo concepto *el análisis de coocurrencias de palabras o análisis co-word*, que estudia la aparición conjunta de dos o más palabras que se encuentran en títulos, abstracts, palabras claves, los claims que se encuentran en las patentes, texto en todo el documento, entre otros, posteriormente se lleva a cabo una matriz de co-ocurrencias y finalmente se visualiza en una gráfica en la que de una forma sintética sea posible observar directamente el grado de vinculación que tienen las palabras, por ejemplo en clusters.

Para mostrar los mapas se lleva a cabo por medio de programas informáticos estadísticos por ejemplo Pajek, Lexximape, Tetralogie entre otros; los cuáles aplican herramientas multidimensionales en donde se ubicarán las palabras de un área de conocimiento.

Otra manera de analizar grandes dominios científicos es por medio de los análisis de citas, entendidos éstos en el sentido en el que los definieron Hjørland y Albrechtsen (1995), es decir, la interacción entre los autores y su papel por medio de las citas.

#### 1.4 Análisis de Citas

El *análisis de citas* que también se conoce como análisis de *cocitaciones o análisis de citas conjuntas* de acuerdo a Callon, Courtial y Penan, 1993:58, permite detectar la aparición simultánea de dos citas que se repiten en un número elevado de artículos. La frecuencia de

la cocitación mide el grado de asociación entre dos documentos o dos autores y permite identificar grupos o *clusters* de artículos próximos que son cocitados frecuentemente lo que permite revelar líneas de investigación de una determinada área. Si por el contrario, dos autores nunca son citados a la vez es una demostración de que existe una falta de relación o lejanía entre sus áreas de especialización. (Escorsa y Maspons, 2001).

Un ejemplo de análisis de citas: Si en varias temas de tesis citan a la vez a Alfred Newman, especialista en bases de datos y Brian Kernighan, especialista en Lenguajes de Programación, se está produciendo la aparición de un nuevo *cluster*, que incorpora elementos procedentes de *Base de Datos y Lenguajes de Programación*.

### **1.4.1 Mapas basados en Análisis de Citas**

En análisis de citas permite detectar la aparición simultánea de dos citas que se repiten en gran número de tesis u otra fuente de información; estas citas con la ayuda de los mapas tecnológicos u otra forma de visualización nos permitirán formar agrupaciones o *clusters*, lo que permitirá ver agrupaciones de documentos que citan a un grupo de fundadores muy próximos (Escorsa y Maspons, 2001)

Sin embargo para llevar acabo lo anterior y verlo reflejado en un mapa que llamaremos mapa de co-citaciones se necesitan seguir una serie de pasos:

Paso 1. Delimitación de la muestra. Consiste en seleccionar los datos de las bases de datos que deben ser incluidos en el área a analizar, lo que puede exigir en algunos casos la consulta previa de expertos.

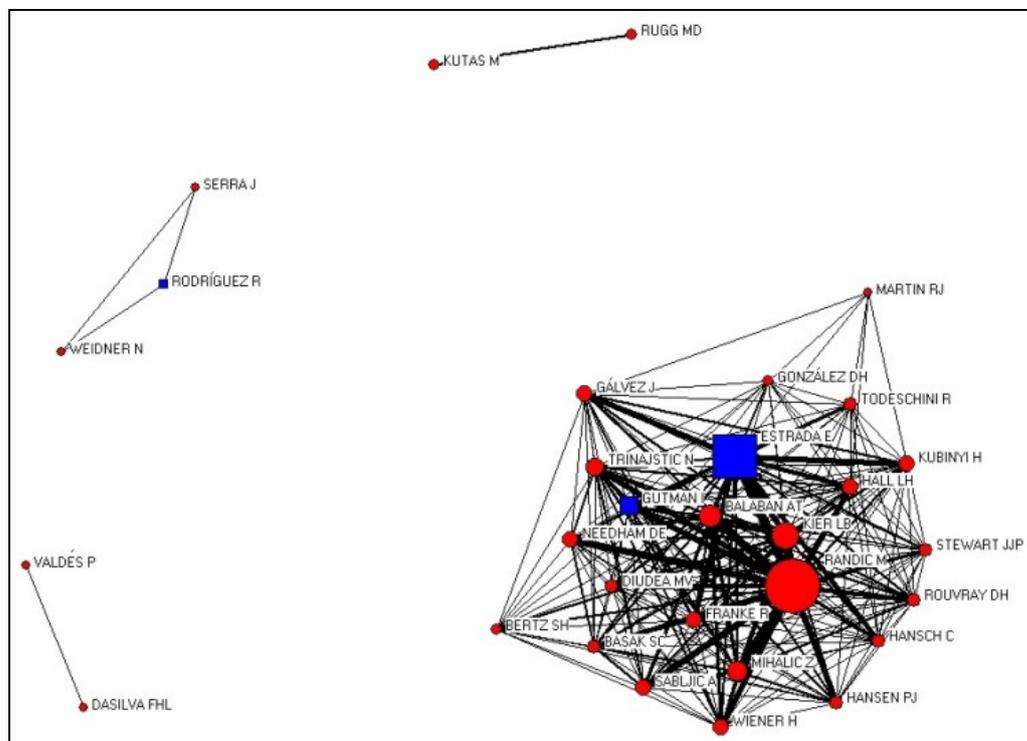
Paso 2. Construcción de la matriz de co-citaciones. Consiste en la elaboración de una lista con todos los autores seleccionados mediante la aplicación de los criterios fijados, que pueden ser autor mas citado, artículo mas citado, lugar de dependencia, entre otros.

Paso 3. Visualización. Consiste en la elaboración del mapa tecnológico que refleje las distancias entre los autores en función de las relaciones de cada uno de ellos con los demás. Su visualización requiere normalmente de un software especializado, algunos requieren licencia.

Cada uno de los grupos obtenidos, que reúne un conjunto homogéneo de citas, forma un *cluster*, o línea de investigación. La visualización obtenida en los mapas tecnológicos es más cualitativa que cuantitativa, es decir no es tan trascendente medir numéricamente las distancias obtenidas como observar más analíticamente lo que está sucediendo con los datos, como la distribución de autores en el espacio y las características de los grupos que aparecen.

Los mapas tecnológicos son herramientas que por medio de figuras geométricas y colores representan relaciones entre objetos como son autores, colaboración entre autores, entre instituciones, libros más citados, artículos mas citados, entre otros. En la obra “Producción de autores cubanos en las revistas sobre ciencias de la computación registradas en el *Journal Citation Report* en el periodo 1990-2005”, Carbonel (2007) hacen un análisis de citas, mostrando por medio de figuras geométricas la red de autores más citados en el JCR y las conexiones entre estos, representando a los autores productores mas citados por medio de un cuadro de color azul; la intensidad de los enlaces entre ellos representa el nivel de citación y el tamaño se relaciona a la cantidad de citas recibidas. Figura 1.4.1

**Figura 1.4. 1 Ejemplo. Mapa de Análisis de Citas**



Fuente: Carbonel. 2007

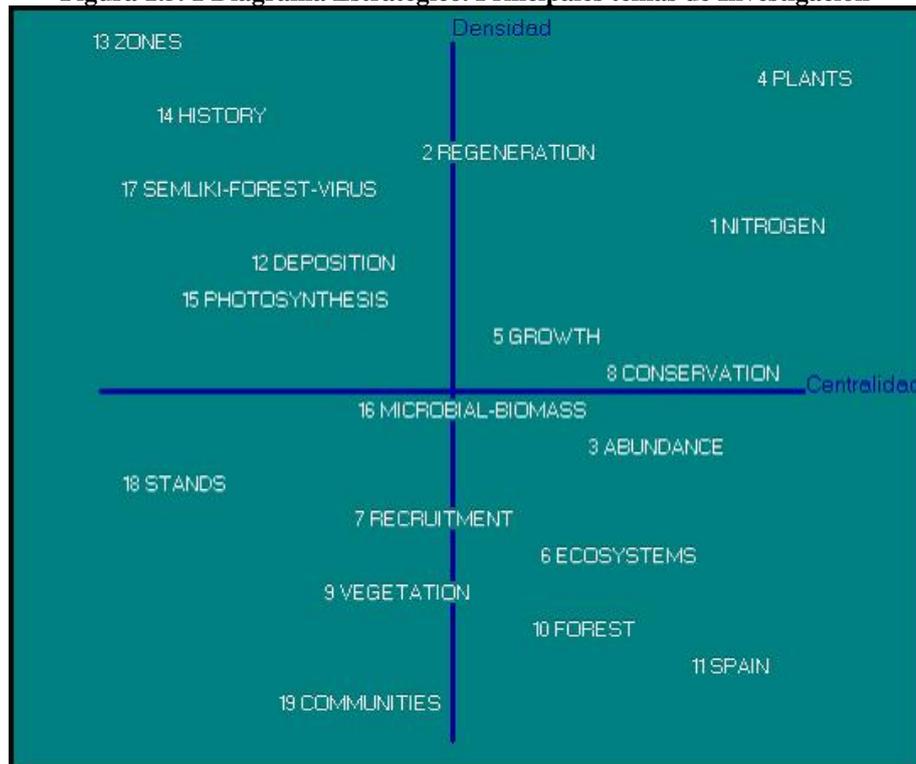
## **1.5 Análisis de Palabras Asociadas**

El método o *análisis de palabras asociadas* que también se conoce como *co-ocurrencia de palabras o análisis co-word* de acuerdo a Callon, Courtial y Penan, 1993:72, consiste en la detección de palabras que caracterizan un tema y en contar la co-aparición de estas, en campos tales como títulos de artículos, palabras clave de artículos, abstracts, códigos de clasificación o bien directamente del texto libre, quien las determina son las personas que lleva a cabo la tarea de introducir la información en las bases de datos y del experto en el área a analizar (Escorsa y Maspons, 2001).

### **1.5.1 Mapas basados en Análisis de Palabras Asociadas**

Los mapas basados en Análisis de Citas permiten detectar, las áreas, clusters o programas de investigación en que se subdivide un campo determinado, pero no permite entrar en el contenido de los documentos. Para conseguirlo es indispensable abordar el análisis de las palabras de títulos de artículos, de tesis, resúmenes, descriptores, entre otros; ya que estas palabras explican el contenido de los propios documentos. Con el Análisis de Palabras Asociadas y los mapas se visualiza la estructura del conocimiento que hay detrás del documento y que lo ha hecho posible (Escorsa, Maspons, 2001).

En la Figura 1.5.1, Díaz (1995) describe el análisis de la producción científica forestal española recogida por ISI, mediante el método de Palabras Asociadas, los resultados se representa en un diagrama estratégico, se obtienen los principales temas de investigación, las subredes de descriptores y las revistas, así como los autores de cada tema.

**Figura 1.5. 1 Diagrama Estratégico. Principales temas de Investigación**

Fuente: Díaz. 1995

Con los avances de la informática, se han desarrollado un considerable número de herramientas estadísticas que son útiles en la cienciometría, agilizando las tareas y multiplicando las posibilidades de análisis. En el siguiente apartado se recopilan algunas y se muestra una breve descripción

### 1.6 Software Informático Aplicado a la Cienciometría y Bibliometría

La información generada y almacenada en el mundo va creciendo cada vez más. Moon (2004) menciona que del 100% de los descubrimientos y logros en Ciencia y Tecnología alrededor del 60% se ha llevado a cabo en los últimos 100 años y de este 60% de descubrimientos y logros e Ciencia y Tecnología el 80% se ha llevado a cabo en los últimos 20 años, como resultado se generan grandes cantidades de información registradas en bases de datos científicas, lo que ha propiciado el desarrollado diferentes tipos de herramientas informáticas con diferentes aplicaciones y características. En la Tabla 1.6 se muestra una breve descripción de algunas herramientas informáticas que son aplicadas a la cienciometría y bibliometría que existen actualmente en el mercado.

**Tabla 1.6 1 Principales Herramientas Estadísticas utilizadas en cienciometría**

No	Nombre de la Herramienta Informática	Descripción
1	BizInt-SmartCharts	Carga Información obtenida de varias Bases de Datos de patentes y de varios Hosts en una hoja electrónica. A continuación permite formatear la información y hacer análisis estadísticos de cada uno de los campos.
2	Delphion.PatentLabII	Permite descargar información de Delphion y hacer diversos análisis de las patentes.
3	Manning & Napier-MapIT	Construido a partir de patentes obtenidas de la USPTO, EPO y WPTO. Permite analizar los principales inventores, patentadores, clasificaciones y citas de patentes, añadir comentarios y almacenar resultados.
4	Delphion-Clustering	Es una función en línea que extrae los términos más relevantes del texto de los resultados de una búsqueda, analiza sus relaciones y los presenta en un mapa.
5	Delphion-Citation Link	Es una función en línea. A partir de una patente conocida, se representa de modo gráfico las relaciones entre patentes citantes y patentes citadas
6	Aurigin-Aureka	Permite hacer búsquedas, análisis de datos de textos, hacer anotaciones y guardar los resultados. Incluye: a) Aureka Themescape: Genera mapas con aspecto geográfico (valles, montañas, desniveles) que permiten entender intuitivamente los conceptos más relevantes, su grado de concentración y grado de relación. b) Aureka Citation Tree: Construye un árbol de citas a partir de una patente seleccionada.
7	Univ. Toulouse-IRIT-Tetralogie	Tetralogie es un paquete muy completo especializado en analizar grandes volúmenes de información científica y de patentes. Su función más elevada es la generación de mapas tecnológicos a partir de análisis estadísticos de co-ocurrencia de términos.
8	Univ. Marsella -CRRM-Dataview/Matrisme	Diseñado sobre todo para analizar información estructurada procedente de Bases de Datos. Permite generar mapas con términos y sus relaciones, tiene un enfoque similar a Tetralogie.
9	VigTech	Permite obtener información de documentos científicos, extracción de meta datos, cálculo de estadísticas descriptivas, análisis de redes sociales, análisis de redes de palabras claves y visualización.
10	HistCite	Permite el análisis de citas y la visualización de gran número de información bibliográfica
11	RefViz	Software que permite la visualización de registros bibliográficos. Permite versión de prueba y es un producto de Thomson Reuters.

Fuente:[http://www.zaintek.net/ebizkaia/Cursos/Curso\\_1/menu/pdf/3.pdf](http://www.zaintek.net/ebizkaia/Cursos/Curso_1/menu/pdf/3.pdf), Fecha de consulta Mayo: 2008

Sin embargo la mayoría de las herramientas mencionadas requieren de una suscripción, por el momento para el análisis de citas se trabajó con el paquete de software llamado

HistCite versión 8.5.26, que se encuentra disponible en <http://www.histcite.com/>, en donde se generan mapas cronológicos de una colección de tópicos resultado de investigaciones, en este caso de la producción científica en lo que respecta a libros, donde se pueden observar los trabajos mas destacados. HistCite versión 8.5.26 también incluye un modulo para detectar y editar errores o variaciones en citas referenciadas así como un analizador de vocabulario el cual genera tanto lista de palabras calificadas y pares de palabras utilizadas en la colección. Algunas de las características de la herramienta HistCite versión 8.5.26 (<http://www.histcite.com/whatcanitdo.htm>):

1. Identifica la literatura clave en un campo de investigación
  - a. Artículos importantes en el desarrollo del tópico
  - b. Autores mas citados y más productivos
  
2. Análisis de productividad en publicaciones y trayectoria dentro de una colección de artículos de investigación. Compara características tales como:
  - a. Países e instituciones a partir de donde los autores publican.
  - b. Autor más productivo y mas citado dentro de un grupo
  - c. Estadísticas de citas para grupos y subgrupos (citas de artículos, numero de autores por artículo)
  
3. Reconstruye la historia y desarrollo de un campo de investigación:
  - a. Artículos altamente citados
  - b. Relación de co-autores importantes
  - c. Línea de tiempo de las publicaciones de los autores
  - d. Muestra por medio de un grafico de historias los principales artículos y línea del tiempo en un campo de investigación.

### **1.7 Aplicaciones de Cienciometría y Bibliometría**

En los siguientes párrafos, se lleva a cabo una descripción de cómo las disciplinas cienciometría y bibliometría se han utilizado para analizar el volumen de información científica.

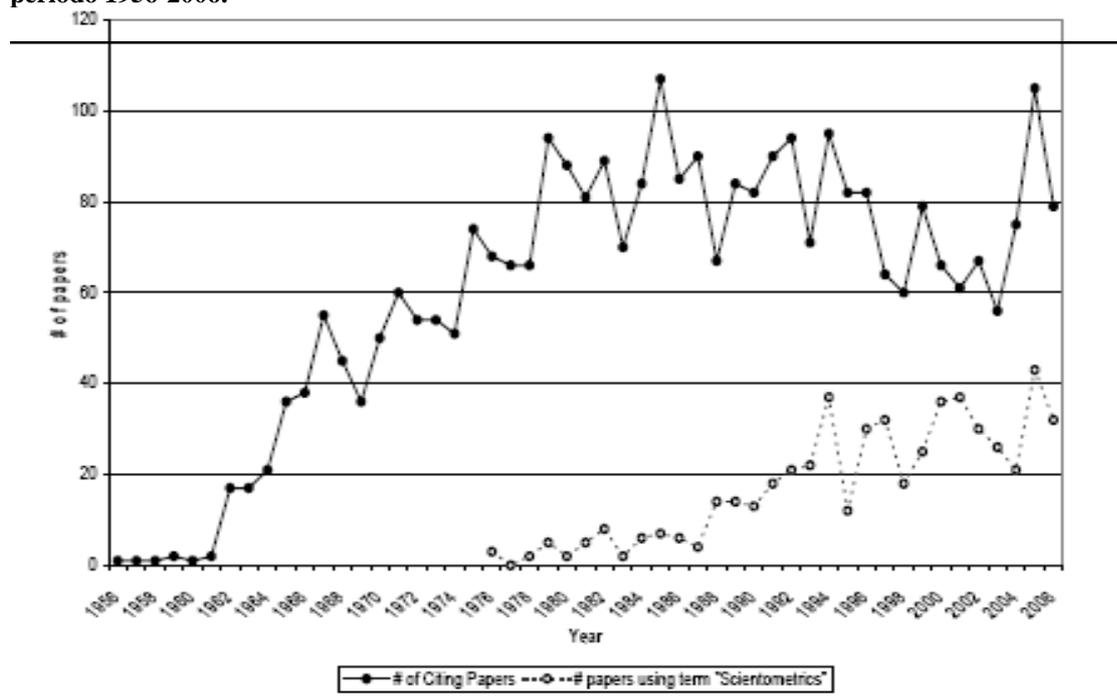
### **1.7.1 Estudio 1: La bibliometría en la Producción de Autores cubanos en las revistas sobre ciencias de la computación registradas en el Journal Citation Reporter en el periodo 1990-2005”**

Esta obra fue publicada en el 2007 por Carbonel S. En el estudio se identifican 95 artículos publicados con la participación de autores cubanos en 44 títulos de revistas, utilizaron indicadores bibliométricos tradicionales, como análisis de redes de cocitación y mapas de co-citación generados por medio de técnicas de escalamiento multidimensional, además se identificaron siete disciplinas con mayor presencia en la producción científica cubana en ciencias de la computación en el Journal Citation Report (ISI 2005), los principales autores, las instituciones, revistas y niveles de colaboración entre los autores. Utilizaron la base de datos documental *Science Citation Index* (ISI 2006), la estrategia de búsqueda utilizó los datos de Journal, author adress, y date of publication.

### **1.7.2 Estudio 2: Un análisis para visualizar la historia de la ciencia**

Este estudio fue publicado en 2007 por Eugene Garfield, siendo fundador del “Instituto de Información Científica (ISI por sus siglas en ingles)”, demuestra que el uso del término *cienciometría* se ha desarrollado a partir de que Price, considerado como el padre de la *cienciometría* publica sus libros “*Science Since Babylon in and Little Science y Big Science*” en 1961 y 1983 respectivamente, lo demuestra usando el software HistCite para visualizar su impacto, así como el posterior impacto en la revista *Scientometrics* con 100, 000 referencias citadas en 3,000 artículos citando a Price; mostrando el análisis de tres graficas, la primera muestra como la frecuencia crece con respecto al tiempo ya sea al citar a Price o utilizar el término *cienciometría* (el orden de la frecuencia de las citas es de 0-120 y los años de 1956-2006). Figura 1.7.2.1.

**Figura 1.7.2. 1 Artículos que citan a Price versus artículos que usan el termino “cienciometría”, en el periodo 1956-2006.**



Fuente: Garfield. 2007

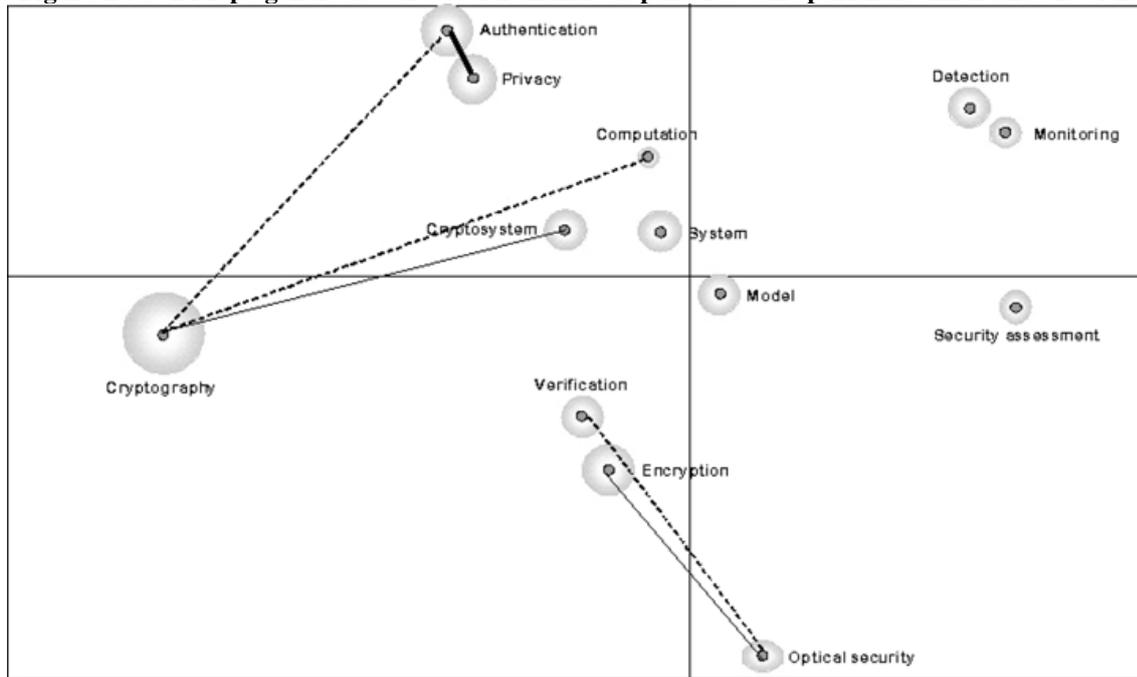
En la segunda gráfica presenta un mapa donde se muestran los trabajos mas citados a partir de 1961-1983, el mapa incluye esferas, de acuerdo al tamaño de esta indica que son los más citados, en este caso lo citan 107 veces. La ultima figura es simplemente mostrar los artículos que se relacionan con la cienciometría a partir de 1917-1969, además del autor e indicando el número de citas por año.

### 1.7.3 Estudio 3: Como identificar campos de investigación emergentes usando la cienciometría: un ejemplo en el campo de la seguridad de la información

Este artículo es publicado por Hyoung (2007), lleva a cabo una investigación que tiene como objetivo el proporcionar indicadores y métodos de visualización para medir las tendencias de investigaciones mas recientes y subrayando aspectos de documentos científicos y tecnológicos para los investigadores y planeadores de políticas usando análisis co-word. El campo de la seguridad de la información es una prospectiva alta en cuanto al valor en el mercado. El análisis co-word es empleado para identificar patrones y tendencias en el campo de la seguridad de la información. Los datos fueron colectados de Science Citation Index y las palabras claves pudieron haber sido extraídas de las palabras claves del autor. Estas palabras claves se extrajeron para después ser estandarizadas. Con la finalidad de rastrear cambios dinámicos en el campo de la seguridad de la información,

se presentan una variedad de tecnologías de mapeo. Los resultados mostraron que el campo de la seguridad de la información ha establecido algunos temas de investigación y además se transforma rápidamente para abarcar nuevos temas, finalmente se concluye que el campo de la Seguridad de la Información tiene algunos temas de investigación ya establecidos y también puede rápidamente transformarse en nuevos temas a tratar con futuro. Figura 1.7.3.1

**Figura 1.7.3. 1 Mapa generado a través del análisis de palabras con apariencia simultanea 1994-1999**



Fuente: Hyoung. 2007

# Capítulo 2

## Aplicación de la Bibliometría

En México no existen estudios sobre la producción científica en el campo de las ciencias de la computación, y en cambio se han realizado estudios de producción científica sobre otras disciplinas como cirugía médica, realizado por (Pineda et al., 2005), en bioinformática por Sánchez (n.d), seguridad de la información (Hyoung, 2007).

Por lo anterior, buscando beneficios similares, se hace un estudio sobre la producción científica de las ciencias de la computación en un centro de investigación de este campo. El estudio es sobre la producción de trabajos de tesis de maestría concluidas.

La metodología a seguir para el desarrollo del presente trabajo, se describe en el anexo 4, esta metodología es el resultado de una recopilación de metodologías de trabajos similares como (Carbonel, 2007, Hyoung, 2007).

Durante el desarrollo de este capítulo, se muestra la información de la base de datos de los generales de las tesis con la que se realizó el estudio bibliométrico, se muestran los resultados de la aplicación de los indicadores bibliométricos de acuerdo a Spinak (1988) que incluyen los indicadores de producción y los indicadores de análisis de citas. En los dos puntos anteriores se llevo a cabo una revisión de los indicadores y se identificaron los que aplican al caso de estudio.

### 2.1 Base de Datos de los Temas de Tesis

Los datos de las 198 tesis pertenecientes al periodo 1996-2007 se recopilaron con información suministrada de la unidad de tecnologías educativas de la biblioteca del centro y con el apoyo de profesores. Los datos se organizaron en hojas de Excel, ejemplo de su contenido se puede apreciar en el Anexo 1.

## 2.2 Principales Indicadores Bibliométricos

Existen diferentes puntos de vista para definir cuáles y cuántos son los indicadores bibliométricos como se describe en (Escorsa y Maspons, 2001), (Spinak E, 1988) y (Vinkler, 1988), Vinkler presenta una tabla comparativa con 46 indicadores simples y compuestos que difieren de los indicadores de la Tabla 1.2 descritos en el capítulo 1 y que son proporcionados por (Escorsa y Maspons, 2001: 94).

El propósito de esta sección es identificar, seleccionar y aplicar los indicadores Bibliométricos a la producción científica compuesta por temas de tesis de maestría.

### 2.2.1. Resultados y discusión de Indicadores de Publicación

La identificación y selección de los indicadores Bibliométricos se llevó a cabo al analizar los indicadores de Spinak (1988), los cuales se dividen en dos grupos: indicadores de publicación e indicadores de citación. Los indicadores de publicación miden la calidad e impacto de las publicaciones científicas según Vinkler (1988). Tabla 2.2.1.

**Tabla 2.2. 1 Indicadores de Publicación**

No	Nombre	Medida
1	Extensión Bibliométrica	Número de artículos por país, por disciplina, etc.
2	Tabla de revistas por especialistas	Primera, segunda, tercera y cuarta clase de revistas
3	Indicador de producción	Numero ponderado de artículos, libros, etc., actividad por número de autores.
4	Índice de actividad	Porcentaje de publicaciones por X / Porcentaje de publicaciones por Y
5	Crecimiento de la documentación	Tasa de crecimiento
6	Distribución de Bradford	Revistas Principales
7	Distribución de Lotka	Productividad de los autores
8	Obsolescencia	Vida media de la publicación
9	Distribución de Zipfs	La frecuencia de aparición de las palabras es inversamente proporcional a la posición que ocupan al clasificar la frecuencia de mayor a menor.
10	Distribución de Waring	Potencial de publicación

**Fuente: Spinak.1988**

Los indicadores para poder ser aplicados se adaptaron a las características del documento base de estudio, la adaptación final de los indicadores se muestra en la Tabla 2.2.2, de aquí en adelante se hará referencia a estos indicadores como indicadores de producción.

**Tabla 2.2. 2 Indicadores de producción en tesis concluidas**

No	Nombre	Medida	Gráfica
1	Producción de Tesis	Número total de tesis	2.3.1.1
2	Indicador de producción	Producción de tesis por laboratorio Producción por director de tesis Producción de tesis con dos directores Red de Colaboración entre Laboratorios Producción de tesis por institución de director externo	2.3.2.1 2.3.2.2 2.3.2.3 2.3.2.4 2.3.2.5
3	Distribución de Lotka	Productividad de los directores de tesis	2.3.3.1
4	Distribución de Zipfs	Uso de vocabulario utilizado en los títulos de las tesis	2.3.4.1

**Fuente: Elaboración propia con base a Spinak.1998**

A continuación se muestran las graficas que corresponden únicamente a la Producción Científica en temas de tesis.

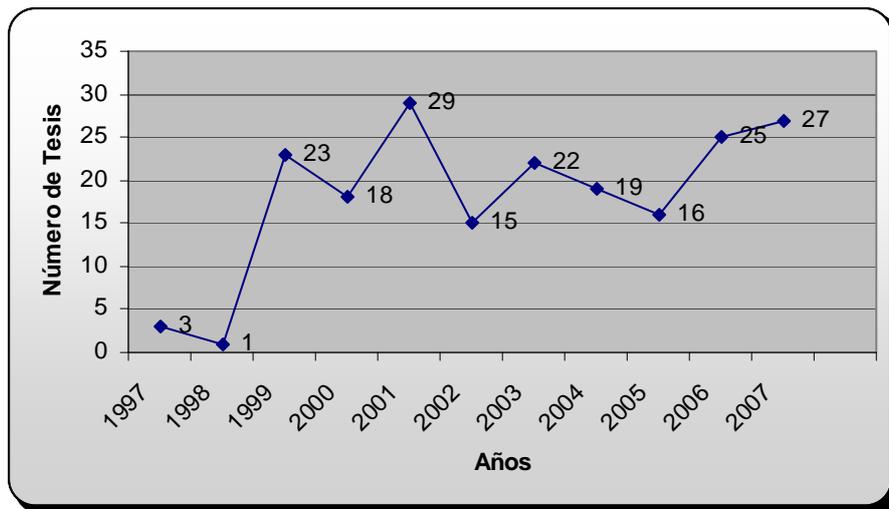
### 2.3 Indicadores de producción científica en tesis concluidas

La producción científica de un investigador o grupo de investigadores puede ser evaluada por varias razones que van desde la cantidad de artículos que publican en revistas reconocidas, patentes, desarrollos tecnológicos, entre otros. El siguiente apartado muestra los indicadores de producción científica utilizados de acuerdo al recurso de información.

#### 2.3.1 Producción de Tesis

##### 2.3.1.1 Numero total de tesis

En la Gráfica 2.3.1.1 se observa 198 tesis, mostrando como se distribuyen en los años del periodo de análisis. Analizando la Gráfica puede notarse un incremento discontinuo con periodos de crecimiento negativo, el pico de la producción científica corresponde al año 2001 con 29 tesis finalizadas, seguida de los años 1999 y 2007 con 23 y 27 trabajos de tesis respectivamente.

**Gráfica 2.3.1. 1 Producción de Tesis sobre ciencias de la computación en el periodo 1996-2007**

Fuente: Elaboración Propia con base en datos del CIC.2007

### 2.3.2 Indicador de Producción.

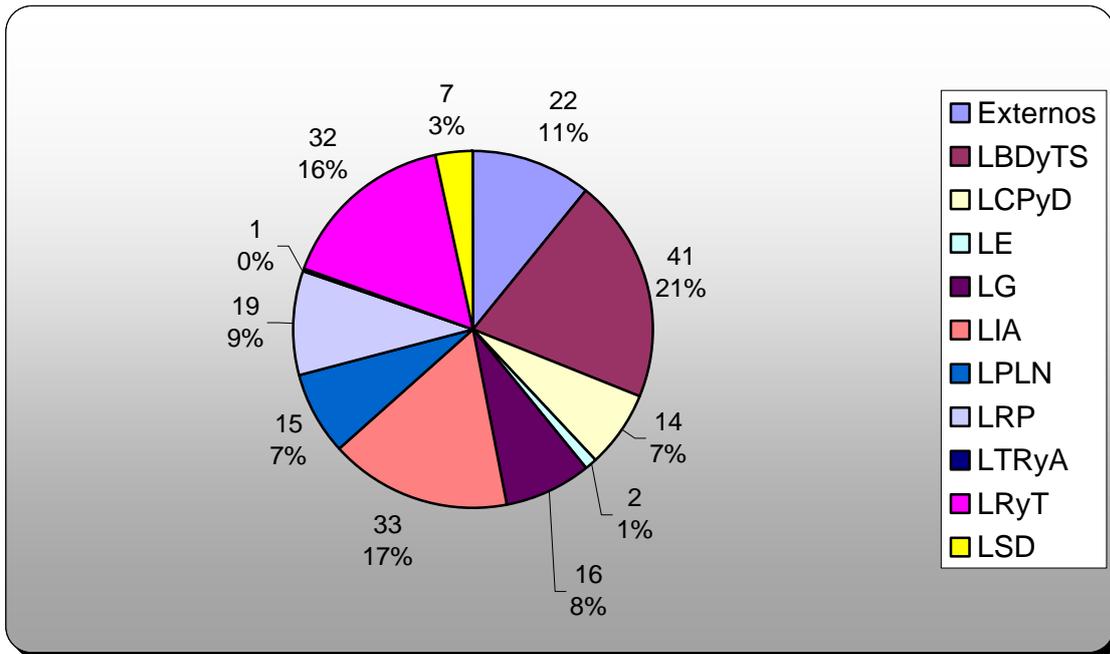
Para analizar la fuente de producción científica dentro del centro de investigación en estudio, se procedió a identificar la forma en que se encuentran agrupados los investigadores responsables de dirigir temas de tesis, por medio de la información que se encuentra publicada en la página oficial del centro de investigación y se identificó que la manera de agrupación es por medio de laboratorios. Los laboratorios identificados son los siguientes:

1. Laboratorio de Base de Datos y Tecnología de Software (LBDyTS)
2. Laboratorio de Electrónica (LE)
3. Laboratorio de Computación Paralela y Distribuida (LCPyD)
4. Laboratorio de Geoprocesamiento (LG)
5. Laboratorio de Inteligencia Artificial (LIA)
6. Laboratorio de Procesamiento de Lenguaje Natural (LPLN)
7. Laboratorio de Reconocimiento de Patrones (LRP)
8. Laboratorio de Redes y Telecomunicaciones (LRyT)
9. Laboratorio de Sistemas Digitales (LSD)
10. Laboratorio de Tiempo Real y Automatización (LTRyA)

### 2.3.2.1 Producción de tesis por laboratorio

Con la agrupación de los directores de tesis por laboratorio, se obtuvo la Gráfica 2.3.2.1, donde se muestra la estadística que nos permite ver como se distribuye *la producción de tesis por laboratorio*, además permite visualizar como se encuentra la participación de un grupo externo de investigadores como directores de tesis.

**Gráfica 2.3.2. 1 Producción de Tesis por Laboratorio**



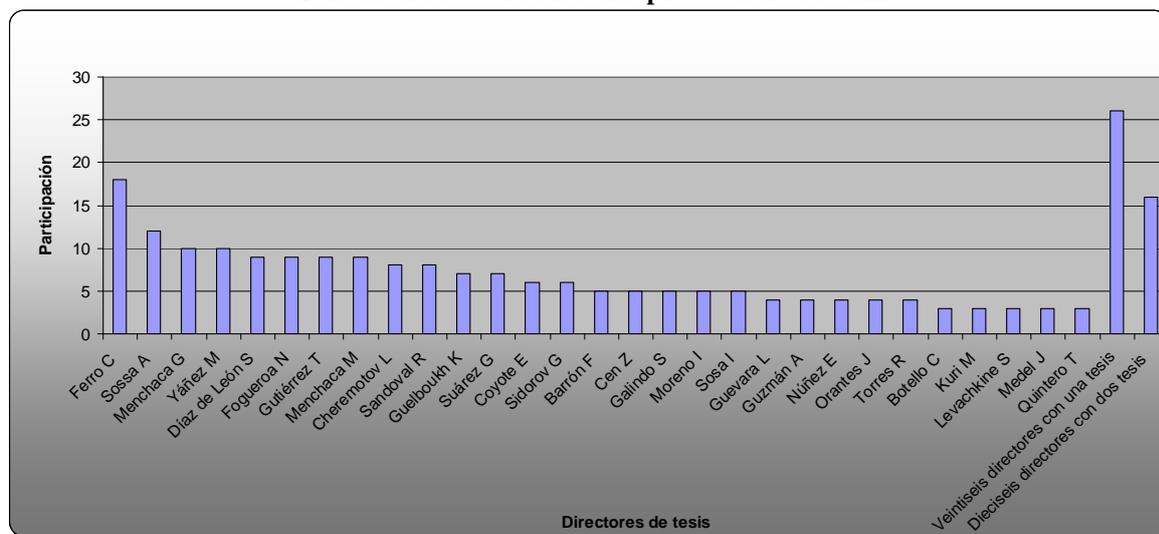
**Fuente: Elaboración Propia con base en datos del CIC.2007**

También de la Gráfica 2.3.2.1, se puede apreciar que los principales laboratorios con mayor producción fueron: Laboratorio de Base de Datos y Tecnología de Software (LBDyTS) con 41 tesis, 21 %; Laboratorio de Inteligencia Artificial (LIA) con 33 tesis, 19%, y Laboratorio de Reconocimiento de Patrones (LRP) con 32 tesis, con el 16 %. La producción conjunta de estos laboratorios representa el 57 %. Los seis laboratorios restantes conforman el 33 % del total; el Laboratorio de Tiempo Real y Automatización (LTRyA) es el que tiene menor producción con una tesis, seguida del Laboratorio Electrónica (LE) con 2 tesis. El 11 % corresponde a investigadores externos que han dirigido temas de tesis.

### 2.3.2.2 Producción por Director de Tesis

En la Gráfica 2.3.2.2 corresponde a la *Producción por Director de tesis*, se observan los directores de tesis internos, donde el director de tesis con un mayor producción resultó ser el investigador *Ferro C.* con 18 trabajos de tesis, seguido por *Sossa A* con 12 y *Menchaca G* con 10. Existe un conjunto de directores de tesis con menos de nueve tesis y más de 3 tesis. Todos los restantes investigadores con direcciones de dos y una tesis se grafican en dos barras.

Gráfica 2.3.2. 2 Productividad por Director de Tesis

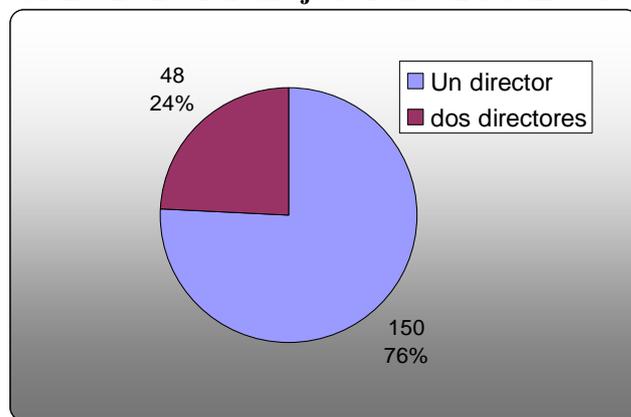


Fuente: Elaboración Propia con base en datos del CIC.2007

### 2.3.2.3 Producción de Tesis con dos Directores

En la Gráfica 2.3.2.3 se observa el porcentaje de *tesis con dos directores*, se puede apreciar que el 24 % de la producción se ha dirigido por dos directores y el 76 % por un director.

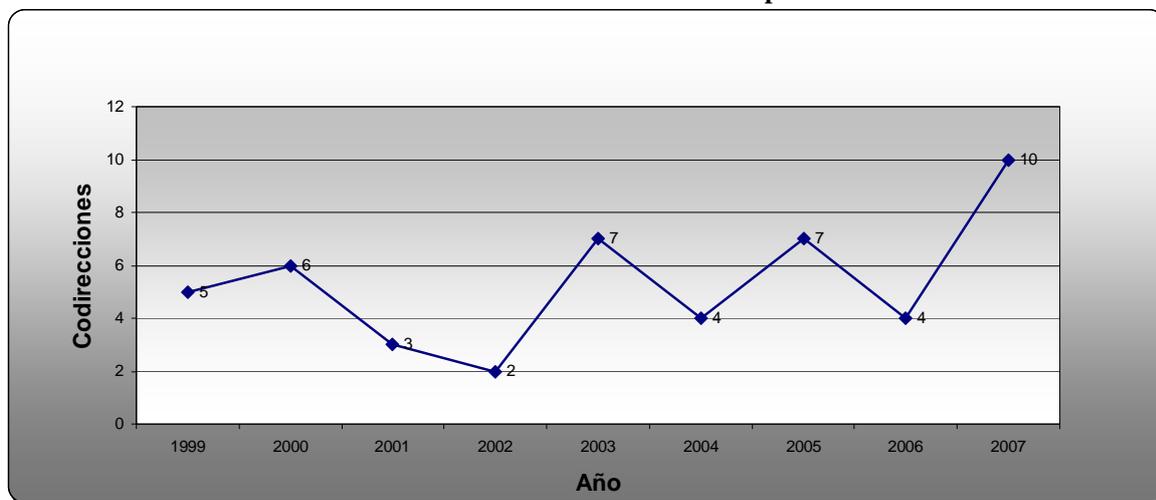
Gráfica 2.3.2. 3 Porcentaje de Tesis con dos directores



Fuente: Elaboración Propia con base en datos del CIC.2007

En la Gráfica 2.3.2.4 se muestra la estadística por año de las 48 tesis codirigidas, y esta permite observar una tendencia de crecimiento en las tesis codirigidas.

**Gráfica 2.3.2. 4 Codirecciones de tesis por año**



Fuente: Elaboración Propia con base en datos del CIC.2010

### 2.3.2.4 Red de Colaboración entre Laboratorios

Es interesante analizar la participación en la dirección de tesis entre los laboratorios del Centro. Para este análisis que se puede ver como una red de colaboración entre laboratorios, esta red se construye con una matriz de tesis dirigidas entre los laboratorios al que pertenecen los investigadores, en la Tabla 2.3.2.1 se aprecia el número de participaciones entre laboratorios.

**Tabla 2.3.2. 1 Matriz de Colaboración en dirección de tesis entre Laboratorios**

	EXTERNO	LBDyTS	LCPyD	LG	LIA	LPLN	LRP	LRyT	LSD	LTRyA	LE
EXTERNO	0	9	2	0	1	1	3	2	1	1	0
LBDyTS	9	5	0	0	2	0	0	0	0	0	0
LCPyD	2	0	0	0	1	0	0	1	1	0	0
LG	0	0	0	3	0	0	0	3	0	0	0
LIA	1	2	1	0	7	0	0	0	0	0	1
LPLN	1	0	0	0	0	2	0	0	0	0	0
LRP	3	0	0	0	0	0	2	0	0	0	0
LRyT	2	0	1	3	0	0	0	0	0	1	0
LSD	1	0	1	0	0	0	0	0	0	0	0
LTRyA	1	0	0	0	0	0	0	1	0	0	0
LE	0	0	0	0	1	0	0	0	0	0	0

Fuente: Elaboración Propia con base en datos del CIC.2007

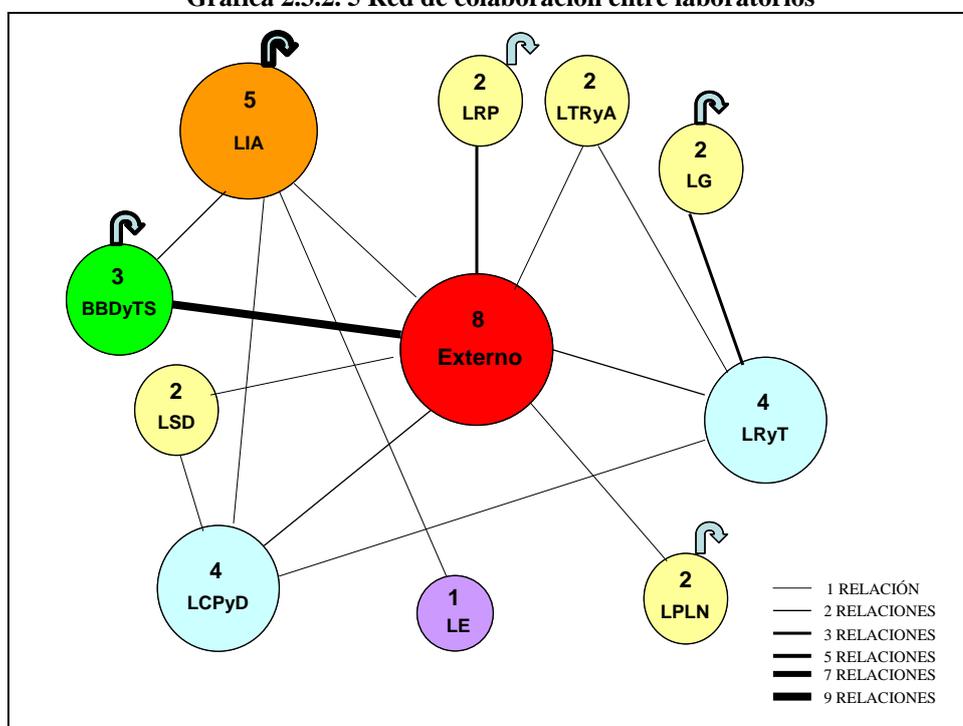
Para la construcción de la tabla anterior se llevo a cabo de la siguiente manera:

1. Se hace un ordenamiento horizontal y vertical por laboratorio

2. Inicialmente las celdas son llenadas con “0s”
3. Cuando exista una relación entre laboratorio columna y laboratorio fila se coloca un 1, en el siguiente ciclo incrementa 1.

La Gráfica 2.3.2.4 muestra la red de *Colaboración entre Laboratorios*, el tamaño del círculo es acorde al número de laboratorios con los que tiene colaboración. Por ejemplo el laboratorio con mayor número de colaboraciones es el Laboratorio de Inteligencia Artificial (LIA), con cinco ligas, las flechas de retorno indican la relación entre ellos mismos. Nuevamente como se observo en la Gráfica 2.3.2.1, el grupo de investigadores externos se destacan y que en este caso se agrupan como un laboratorio para poder ser graficado.

**Gráfica 2.3.2. 5 Red de colaboración entre laboratorios**

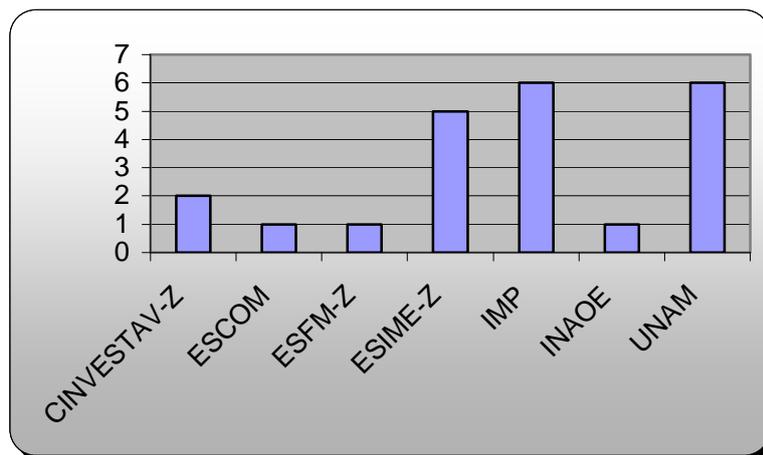


Fuente: Elaboración Propia con base en datos del CIC.2007

### 2.3.2.5 Producción de Tesis por Institución de Director Externo

El nodo de la Grafica 2.3.2.4 representado con círculo rojo identifica la procedencia de investigadores externos, siendo las siguientes instituciones con las que el centro tiene conexiones CINVESTAV-Z, ESCOM, ESFM-Z, ESIME-Z, IMP, INAOE y UNAM. Gráfica 2.3.2.5.

Gráfica 2.3.2. 6 Procedencia de Investigadores Externos



Fuente: Elaboración Propia con base en datos del CIC.2007

### 2.3.3 ley de Lotka

En este apartado se lleva a cabo un breve análisis de la Ley de Lotka que expresa la distribución de los autores de acuerdo a su productividad. Se inicia a partir de la gráfica 2.3.2.2 mostrada en la sección *Producción por Director de tesis* en la que se puede ver la estadística del número de tesis dirigidas por investigador. Las barras (izquierda-derecha) indican la productividad por investigador en las guías de tesis.

La estadística de la Figura 2.3.2.2 se aproxima a la conclusión de la Ley de Lotka, que comenta que en un campo, existen un pequeño grupo de investigadores muy productivos y el restante que son mucho menos productivos (Gorbea, 2005:3).

Para nuestro análisis partiremos de la siguiente fórmula de esta ley:

$$A_n = \frac{A_1}{n^2}$$

donde:

$A_n$  = Es el número de autores con  $n$  firmas.

$A_1$  = Número de autores con 1 firma.

$n^2$  = Es el número de firmas al cuadrado

El análisis aquí presentado se tomó de [http://www.ugr.es/~rruizb/cognosfera/sala\\_de\\_estudio/bibliometria/lotka.ppt](http://www.ugr.es/~rruizb/cognosfera/sala_de_estudio/bibliometria/lotka.ppt), existen otros

métodos como el de (Gorbea, 2005), pero por sencillez en su explicación de la fuente se tomara el ya indicado. Se muestran las frecuencias por dirección de tesis. Tabla 2.3.3.1.

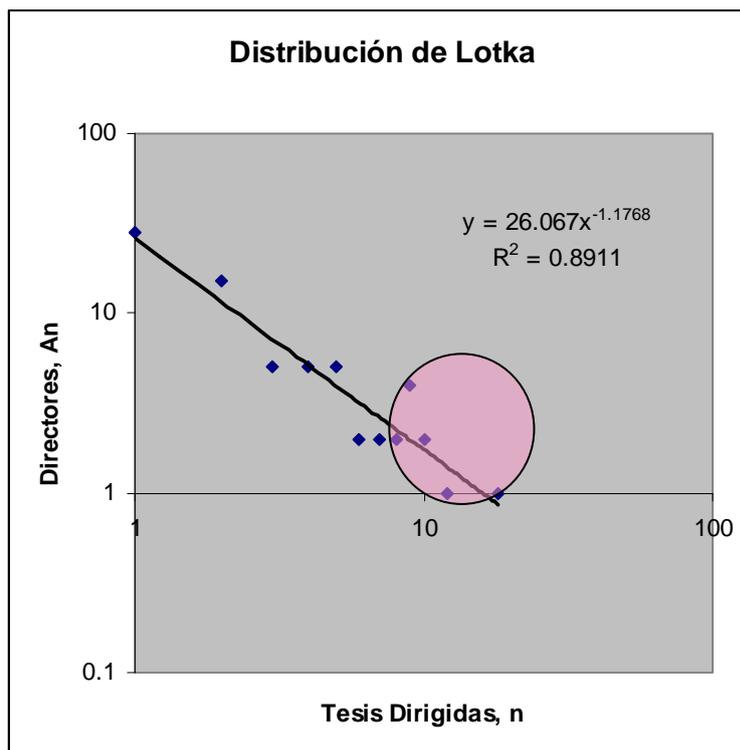
**Tabla 2.3.3. 1 Distribución de frecuencias por direcciones de tesis**

Tesis Dirigidas, n	Directores, An
1	26
2	16
3	5
4	5
5	5
6	2
7	2
8	2
9	4
10	2
12	1
18	1

**Fuente: Elaboración Propia con base en datos del CIC.2007**

Ruiz (nd) propuso que se consideren como más productivos aquellos que representan la raíz cuadrada de todos los autores, en este caso directores de tesis que son 71, si sumamos la segunda columna de la Tabla 2.3.3.1. Así, los directores de tesis más productivos son:  
 $8.5 = \sqrt{71}$ . Grafica 2.3.3.1

Gráfica 2.3.3. 1 Distribución de Lotka



Fuente: Elaboración Propia con base en datos del CIC.2007

Se lleva a cabo un proceso más analítico por el método de los mínimos cuadrados, se esperaba llegar a los mismos resultados, sin embargo los resultados fueron diferentes. Parte de los cálculos se muestran en la Tabla 2.3.3.2

Tabla 2.3.3. 2 Distribución de frecuencias por direcciones de tesis

Tesis, n	Directores, An	X (log(n))	Y (log(D))	XY	X*X	Autores Acumulados (de mas productivos a menos productivos)
1	28	0.000000	1.447158	0.000000	0.000000	72
2	15	0.301030	1.176091	0.354039	0.090619	44
3	5	0.477121	0.698970	0.333493	0.227645	29
4	5	0.602060	0.698970	0.420822	0.362476	24
5	5	0.698970	0.698970	0.488559	0.488559	19
6	2	0.778151	0.301030	0.234247	0.605519	14
7	2	0.845098	0.301030	0.254400	0.714191	12
8	2	0.903090	0.301030	0.271857	0.815572	10
9	4	0.954243	0.602060	0.574511	0.910579	8
10	2	1.000000	0.301030	0.301030	1.000000	4
12	1	1.079181	0.000000	0.000000	1.164632	2
18	1	1.255273	0.000000	0.000000	1.575709	1
<b>Sumas</b>	<b>72</b>	<b>8.894217</b>	<b>6.526339</b>	<b>58.046676</b>	<b>79.107092</b>	

Fuente: Elaboración Propia con base en datos del CIC.2007

$$b = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2} = \frac{12 \times 58.046676 - 8.894217 \times 6.526339}{12 \times 79.107092 - (8.894217)^2} = 0.73377335$$

$$a = \frac{\sum y_i - b \sum x_i}{n} = \frac{6.526339 - 0.73377335 \times 8.894217}{12} = -3.36431E-08 \xrightarrow{\text{Sededucque}} A_1 = 10^{-3.36431E-08} = 0.999999925340 \text{ autores} > 28$$

Concluyo que algún error cometo, pues no conseguí obtener los mismos resultados con la aplicación del método de mínimos cuadrados.

### 2.3.4 Ley de Zipf

La llamada Ley empírica de Zipf, comenta sobre el uso y frecuencia de las palabras que se utilizan en el texto de un documento. Indica que hay un pequeño número de palabras que son utilizadas con mucha frecuencia, mientras que un gran número de palabras son poco empleadas. [http://es.wikipedia.org/wiki/Ley\\_de\\_Zipf](http://es.wikipedia.org/wiki/Ley_de_Zipf), consulta realizada el 15 de julio del 2010.

La afirmación relaciona las palabras en su posición (R) de la lista de frecuencias (F) con una constante (k), que se expresa en forma matemática como sigue:

$$F = \frac{k}{R}$$

Para verificar la *Ley de Zipf*, se utilizaron las palabras que aparecen en los títulos de 202 tesis que se registraron en una colección con el producto de software llamado *HistCite* 8.5.26, ver Figura 2.3.4.1. La verificación servirá para observar y dar una primera opinión sobre que tópicos de la computación tratan los desarrollos de tesis.

Figura 2.3.4. 1 Colección de Tesis en HistCite

#	Date / Author / Journal	LCS	GCS	LCR	CR
<b>1997</b>					
1	1 Garcia MM Sistema reporteador para informacion del servicio telefonico medido (SRSISTEM) CIC-IPN. 1997 MAY 09;	0		0	0
2	2 Monroy SE Sistema generador automatico de reportes en "C" para acceso a SQL CIC-IPN. 1997 OCT 10;	0		0	0
3	3 German IA Sistema computarizado de automatizacion de calculos para el diseño de secadores por aspersion (DI.S.A.D.A.) CIC-IPN. 1997 DEC 12;	0		0	0
<b>1998</b>					
4	4 Hernandez GC Sistema grafico para edicion de pozo CIC-IPN. 1998 NOV 03;	0		0	0
<b>1999</b>					
5	5 Paredes VC Manipulador topografico de la Republica Mexicana con despliegue grafico CIC-IPN. 1999 ABR 16;	0		0	0
6	6 Chac KM Herramientas para el agrupamiento de objetos simbolicos CIC-IPN. 1999 MAR 17;	0		0	0
7	7 Gonzalez GE Herramientas para el proceso e interpretacion de datos gravimetricos y magnetometricos empleando una interfaz grafica CIC-IPN. 1999 MAR 19;	0		0	0
8	8 Taddei ZE Un algoritmo genetico para la optimizacion de estructuras metalicas CIC-IPN. 1999 MAY 11;	0		0	0

Fuente: Elaboración Propia con base en datos del CIC, utilizando la herramienta HistCite.2007

El software *HistCite 8.5.26* ayuda a extraer las palabras de los títulos y calcula su frecuencia, además de que permite ordenarlas en diferentes formatos (frecuencia, orden alfabético, entre otros), e indicando también la fuente de las palabras. Como resultado se obtiene un total de 794 palabras en los títulos de 202 tesis. Ver anexo 5 con la lista completa de las palabras de los títulos.

La Tabla 2.3.4.1 muestra las primeras 200 palabras, una primera afirmación sobre las palabras en los títulos, es que son palabras del vocabulario de computación. En los renglones en blanco de la Tabla 2.3.4.1 se omitieron las palabras vacías (Luhn, 1958).

**Tabla 2.3.4. 1 Vocabulario en los títulos de las Tesis**

Rango	Frecuencia	Palabras/Descriptor									
1	103		51	5	ASIGNACION	101	3	CONTEO	151	2	CONOCIMIENTO
2	51	SISTEMA	52	5	BETA	102	3	DICCIONARIO	152	2	CONTEXTO
3	26	DATOS	53	5	DETECCION	103	3	EDIFICIOS	153	2	CONVERSION
4	22		54	5		104	3	ESPACIALES	154	2	DECISIONES
5	21	ANALISIS	55	5	ESTRUCTURAS	105	3	ESPACIO	155	2	DIFERENTES
6	19		56	5	MULTIAGENTE	106	3	GENERADOR	156	2	DIGITAL
7	18		57	5	MUNDOS	107	3	IDENTIFICACION	157	2	DIGITALES
8	17	MODELO	58	5	RECONOCIMIENTO	108	3	INDUSTRIALES	158	2	DISTRIBUIDOS
9	15	CONTROL	59	5	RECUPERACION	109	3	INTERFAZ	159	2	DOCUMENTOS
10	15	DESARROLLO	60	5	TECNICAS	110	3	LENGUAJE	160	2	DSP56827
11	14	SISTEMAS	61	5	VIRTUAL	111	3	MANEJO	161	2	ELECTRONICA
12	13	DISEÑO	62	4	AGENTES	112	3	MEDIO	162	2	EMPLEANDO
13	12	OBJETOS	63	4	ALGORITMO	113	3	MOMENTOS	163	2	ENTRADA
14	12		64	4	APLICACIONES	114	3	MONITOREO	164	2	ESCRITORIO
15	12	USANDO	65	4	AUTOMATIZACION	115	3	MOVIL	165	2	ESPACIAL
16	11	REDES	66	4	BASADA	116	3	NEURONALES	166	2	ESTADISTICO
17	11	TIEMPO	67	4	CALIDAD	117	3	OPTIMIZACION	167	2	EXPERTO
18	10	COMPONENTES	68	4	COMPORTAMIENTO	118	3	PLANTAS	168	2	EXTENSION
19	10	SOFTWARE	69	4	CONTROLADOR	119	3	PRODUCCION	169	2	EXTRACCION
20	9	ESTUDIO	70	4	DISTRIBUIDAS	120	3	PROGRAMACION	170	2	FLUJO
21	9	EVALUACION	71	4	DISTRIBUIDO	121	3	ROBOT	171	2	FUENTES
22	9	IMAGENES	72	4	EDUCACION	122	3	SEMANTICAS	172	2	GENERACION
23	9	INFORMACION	73	4	GEOMETRICOS	123	3	SERIES	173	2	GENERALIZACION
24	9		74	4	GESTION	124	3	SERVIDOR	174	2	GENETICO
25	8		75	4	INFRAESTRUCTURA	125	3	SIMILITUD	175	2	GRAFICA
26	8	ESPAÑOL	76	4	INTERNET	126	3	TEXTOS	176	2	GRAFICO
27	8	SIMULACION	77	4	JAVA	127	3	TRAFICO	177	2	HARDWARE
28	8	WEB	78	4	MEDIANTE	128	3		178	2	HERRAMIENTAS
29	7	AMBIENTE	79	4	MEDICION	129	3	UBICUO	179	2	IBM
30	7	APLICACION	80	4	PALABRAS	130	3	VEHICULAR	180	2	IMAGEN
31	7	ASOCIATIVAS	81	4	PLATAFORMA	131	3	VISUAL	181	2	IMPRESAS
32	7	BASES	82	4	RECURSOS	132	2	ACCESO	182	2	INGENIERIA
33	7	HERRAMIENTA	83	4	SERVICIO	133	2	AGREGADOS	183	2	INTEGRACION
34	7		84	4	SERVICIOS	134	2	AGRUPAMIENTO	184	2	INTEGRAL
35	7	PATRONES	85	4	TECNICA	135	2	ALTO	185	2	INTERPRETE
36	7	REAL	86	4	VRML	136	2	APLICADOS	186	2	INTRANET
37	7	VIRTUALES	87	3	ADMINISTRADOR	137	2	ASFALTENICOS	187	2	JINI
38	6	APRENDIZAJE	88	3	ADQUISICION	138	2	AUTOMATAS	188	2	KHEPERA
39	6	AUTOMATICA	89	3	ALGORITMOS	139	2	BAJO	189	2	LERNMATRIX
40	6	AUTOMATICO	90	3	APOYO	140	2	BASADOS	190	2	LINEALES
41	6	BASE	91	3	BIBLIOTECA	141	2	CALCULOS	191	2	LOCAL
42	6	IMPLEMENTACION	92	3	CALCULO	142	2	CAPACIDAD	192	2	MAPAS
43	6	MEMORIAS	93	3	CLASIFICADOR	143	2	CARACTERIZACION	193	2	MEXICANA
44	6	METODOLOGIA	94	3	COLOCACIONES	144	2	CELULAR	194	2	MINERIA
45	6		95	3		145	2	COLABORATIVOS	195	2	MODELADO
46	6	RED	96	3	COMPARATIVO	146	2	COLOR	196	2	MODELOS
47	6	SOBRE	97	3	COMPLEJO	147	2	COMPILACION	197	2	MODULO
48	6	UTILIZANDO	98	3	COMPUTACIONAL	148	2	COMPRESION	198	2	MORFOLOGICA
49	5	ALFA	99	3	COMUNICACION	149	2	COMPUTADORA	199	2	MOVILES
50	5	ARQUITECTURA	100	3	CONSTRUCCION	150	2	COMPUTADORAS	200	2	ORIENTADOS

Fuente: Elaboración Propia con base en datos del CIC.2007

Tabla 2.3.4. 2 Palabras con mayor frecuencia en los títulos de las tesis

Rango	Frecuencia	Palabras/Descriptor
1	103	PARA
2	61	SISTEMA
3	28	DATOS
4	22	DEL
5	21	ANÁLISIS
6	19	UNA
7	18	CON
8	17	MODELO
9	15	CONTROL
10	15	DESARROLLO
11	14	SISTEMAS
12	13	DISEÑO
13	12	OBJETOS
14	12	POR
15	12	USANDO
16	11	REDES
17	11	TIEMPO
18	10	COMPONENTES
19	10	SOFTWARE
20	9	ESTUDIO
-	-	-
-	-	-
-	-	-
790	1	WAVELETS
791	1	WBE
792	1	XML
793	1	YACIMIENTOS
794	1	ZONAS

Fuente: Elaboración Propia con base en datos del CIC, utilizando la herramienta HistCite.2007

La Tabla 2.3.4.2 muestra un resumen de las 794 palabras y se utiliza para ilustrar un reordenamiento de las frecuencias de las palabras, es decir se identifican los renglones de las palabras con la misma frecuencia, posteriormente se deja solo la palabra con la última frecuencia, obteniéndose así 23 palabras. Tabla 2.3.4.3.

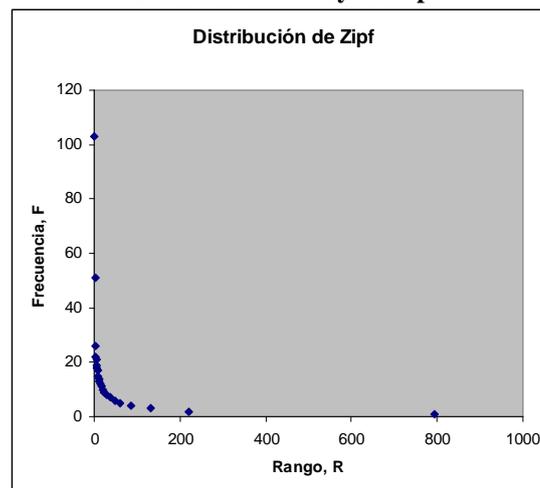
Tabla 2.3.4. 3 Vocabulario ordenado por frecuencias

Rango	Frecuencia	Palabras/Descriptor
1	103	PARA
2	51	SISTEMA
3	26	DATOS
4	22	DEL
5	21	ANALISIS
6	19	UNA
7	18	CON
8	17	MODELO
10	15	DESARROLLO
11	14	SISTEMAS
12	13	DISEÑO
15	12	USANDO
17	11	TIEMPO
19	10	SOFTWARE
24	9	LOS
28	8	WEB
37	7	VIRTUALES
46	6	UTILIZANDO
61	5	VIRTUAL
86	4	VRML
131	3	VISUAL
221	2	VISUALIZACION
794	1	ZONAS

Fuente: Elaboración Propia con base en datos del CIC.2007

En la Tabla 2.3.4.1, se observa que hay una aproximación a la distribución de Zipf, su frecuencia con respecto al rango es aproximada, por ejemplo la frecuencia de la segunda palabra *sistema* es 1/2 de la frecuencia de la primer palabra *para*, la siguiente palabra *datos* es 1/3 de la frecuencia de la primer palabra y así sucesivamente. Gráfica 2.3.4.1.

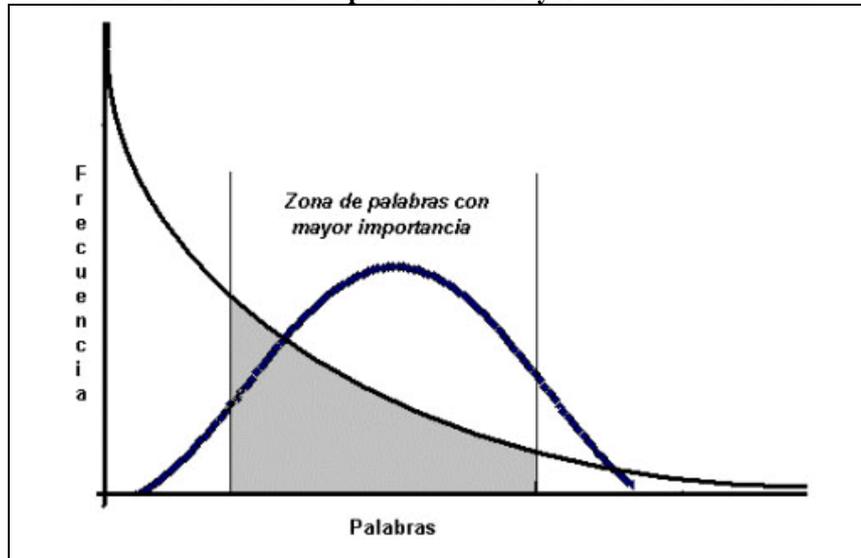
Gráfica 2.3.4. 1 Ley de Zipf



Fuente: Elaboración Propia con base en datos del CIC.2007

(Luhn, 1958), sugirió que las palabras que describen de mejor forma el contenido de un documento se encuentran en una zona de entre las más frecuentes y las de baja frecuencia, en esta zona se encuentran las palabras que mejor describen el contenido semántico de un documento. Gráfica 2.3.4.2

Gráfica 2.3.4. 2 Zona de palabras con mayor contenido semántico



Fuente: Luhn.1958

Esta investigación para su desarrollo en títulos, resúmenes o el contenido de las tesis se deja pendiente para un trabajo a futuro.

## 2.4 Resultados y Discusión de indicadores de citación

En esta sección se llevarán a cabo los indicadores de análisis de citas, que miden la cantidad e impacto de las vinculaciones o relaciones entre las publicaciones científicas y se hará de acuerdo a Spinak (1988), que las resume de una publicación de Vinkler (1988). Tabla 2.4.1.

**Tabla 2.4. 1 Indicadores de citación**

No	Técnica	Medida
1	Activity index	Índice de actividad
2	Affinity index	Índice de afinidad
3	Attractivity index	Índice de atracción
4	Bibliographic coupling	Apareo bibliográfico
5	Citation análisis	Análisis de citaciones
6	Cocitation analysis	Análisis de cocitaciones
7	Consumption factor	Factor de consumo
8	Diversity index	Índice de diversidad
9	Echo factor	Factor de echo
10	Immediacy index	Índice de inmediatez
11	Impact factor	Factor de impacto
12	Impact index	Índice de impacto
13	Insularity index	Índice de aislamiento
14	Openess index	Índice de apertura
15	Popularity factor	Factor de popularidad
16	Self-citation rate	Índice de autocitación

**Fuente:Spinak. 1988**

Por los cálculos requeridos de estos indicadores se busco una herramienta informática de apoyo, localizándose disponible la herramienta HistCite versión 8.5.26, la cuál se facilita y se instala a prueba durante un periodo de 30 días y se encuentra disponible en <http://www.histcite.com/>. Algunas de las herramientas que se menciona en el primer capítulo no se pudieron acceder, ya que se requieren de una licencia y estas de un costo económico.

Por las características de la captura de datos y visualización de la información en la herramienta Informática HistCite versión 8.5.26, ligado a la excesiva mala codificación de las referencias bibliográficas que se tienen en los documentos de tesis, se decidió solamente trabajar con los libros de texto que se citan en estos documentos de tesis a analizar. La codificación errónea obliga a buscar las referencias bibliográficas correctas en páginas de Internet, proceso que se describe en el Anexo 2. Por lo anterior la Tabla 2.4.1 se adaptó como se muestra en la Tabla 2.4.2.

Tabla 2.4. 2 Análisis para aplicar los indicadores de citación.

No	Medida	Definición	Gráfica
1	Índice de afinidad	Puede representar la relación entre dos tecnologías u otros dos eventos que se desee medir. Arriba del 100 % es afín.	Sin grafica
2	Índice de atracción	Es el porcentaje de un número de personas que persiguen un objetivo versus los que logran el objetivo.	Sin grafica
3	Apareo bibliográfico	El apareo bibliográfico se produce cuando dos documentos hacen referencia por lo menos a una misma publicación.	2.4.1 2.4.2 2.4.3
4	Análisis de citaciones	Para Sung y Jung (2008), las clasifica en seis áreas: <b>1. La naturaleza de las citas:</b> Indicar área en que pertenecen <b>2. Autocitación:</b> Se refiere a la proporción de citas que recibe un autor, y que son originadas por el mismo autor. <b>Este índice no aplica, por que no tenemos las citas externas tanto a libros como a temas de tesis.</b> <b>3. Artículos altamente citados:</b> En nuestro caso de estudio se hizo con libros <b>4. Análisis sobre campos especiales:</b> cualquier otro dato de interés por el cual se desee analizar <b>5. Usando aspectos:</b> Se relaciona con aspectos sociales y culturales de como se llevan a cabo las citas. <b>Este índice queda fuera de este estudio.</b> <b>6. Cocitación:</b> Si un artículo de un autor A y otro de un autor B, son citados conjuntamente por un tercer artículo.	2.4.4 2.4.5 2.4.6 2.4.7 2.4.8- 2.4.12 ----- 2.4.13- 2.4.15
5	Análisis de cocitaciones	Ver indicador 5 de esta tabla	-----
6	Índice de impacto	Es el cociente entre el número de citas recibidas y el número de trabajos publicados.	2.4.16
7	Índice de autocitación	Ver indicador 5 de esta tabla	-----

Fuente: Elaboración Propia a partir de Spinak.1988

Los indicadores de Actividad, Afinidad, Factor de Consumo, Índice de Diversidad, Factor de echo, Índice de Inmediatez, Factor de Impacto, Índice de Aislamiento, Índice de Apertura, Factor de Popularidad no se tomaron en consideración, debido a que no aplican al recurso de información de este estudio.

Para ver el detalle teórico y cálculo de los indicadores de citación que se muestran en la Tabla 2.4.2, ver Anexo 3. A continuación se muestra lo que fue posible aplicar de los indicadores de citación.

### 2.4.1 Apareo Bibliográfico

Para el cálculo de este índice, se utilizó la herramienta HistCite versión 8.5.26, se listaron todas las referencias bibliográficas (libros) y se seleccionó aquella que tuviera dos citas. Figura 2.4.1.

2.4. 1 Libro con dos citas

File Analyses View Tools Help		HistCite™			
Colección de 202 Tesis y Citas Bibl, 1996 - 2007		List Totals: LCS 188 List span: 1933 - 2006			
List of 186 Records for Document Type Book					
	<b>Software</b> ADDISON WESLEY PROFESSIONAL COMPUTING SERIES. 1995;				
50	50 Motwani R <b>Randomized Algorithms</b> CAMBRIDGE UNIVERSITY PRESS. 1995;	1		0	0
#	Date / Author / Journal	LCS	GCS	LCR	CR
51	51 Rosen KH <b>Discrete Mathematics and Its Applications</b> MCGRAW HILL. 1995;	1		0	0
52	52 Signore R <b>The ODBC Solution</b> MCGRAW HILL. 1995;	2		0	0
53	53 Geiger K <b>INSIDE ODBC</b> MICROSOFT PRESS. 1995;	1		0	0
54	54 Hassoun MH <b>Fundamentals of Artificial Neural Networks</b> MIT PRESS. 1995;	1		0	0
55	55 KENDALL KE, KENDALL JE <b>Systems Analysis and Design</b> PRENTICE HALL. 1995;	1		0	0

Fuente: Elaboración Propia con base en datos del CIC, utilizando la herramienta HistCite.2007

Haciendo uso de la liga a las dos citas, se pudieron ver las dos tesis que referencia a este libro. Figuras 2.4.2. y 2.4.3

2.4. 2 Dos Tesis citan a un mismo libro

#	Date / Author / Journal	LCS	GCS	LCR	CR
<b>2002</b>					
1	259 Trejo SG Modulo XML para acceder al sistema administrador de bases de datos SQLmx a traves de Internet CIC-IPN. 2002 JUN 14;	0		18	18
<b>2005</b>					
2	321 Gonzalez EL Interfaz para la publicacion de las operaciones de consulta en un sistema administrador de bases de datos (SABD) como servicio web CIC-IPN. 2005 MAY 04;	0		49	49

Fuente: Elaboración Propia con base en datos del CIC, utilizando la herramienta HistCite.2007

En la Figura 2.4.3 se observa que la tesis 259 y la tesis 321, hacen referencia a la publicación de Signore R.

2.4. 3 Apareo Bibliográfico

Fuente: Elaboración Propia con base en datos del CIC, utilizando la herramienta HistCite.2007

## 2.4.2 Análisis de citaciones.

Se estudian 908 citaciones correspondientes a 70 tesis, lo cual arroja un coeficiente de 12.97 referencias por tesis con un rango que oscila entre 1 y 54. Tabla 2.4.3

**Tabla 2.4. 3 Frecuencias de citas**

No	AU	PY	LCR	No	AU	PY	LCR
1	PALOMINIO, RM	1999	9	36	MEDINA, MV	2004	26
2	MENDOZA, AL	1999	29	37	MENCHACA, MR	2005	32
3	MARTINEZ, CM	1999	23	38	MEDINA, AJ	2005	16
4	MANJARREZ, SJ	2000	11	39	SANCHEZ, CM	2005	1
5	PIÑON, MG	2000	6	40	SALGADO, RJ	2005	41
6	SOTELO, RR	2000	11	41	LOPEZ, RG	2005	37
7	SILVA, FJ	2001	12	42	GONZALEZ, EL	2005	49
8	AVILA, MJ	2001	16	43	OLIVAS, ZO	2006	27
9	VARGAS DE BASTE	2001	18	44	ANGELES, YM	2006	6
10	GAMA, ML	2001	8	45	CUEVAS, CW	2006	14
11	OLGUIN, SJ	2001	23	46	DE LA CRUZ, SJ	2006	54
12	AMOROS, ZE	2001	5	47	LANDASSURI, MV	2006	6
13	GERMAN, SE	2002	9	48	FARARONI, RI	2006	29
14	LARIOS, OM	2002	31	49	TORRES, RS	2006	29
15	REYES, DL	2002	8	50	RODRIGUEZ, LJ	2006	1
16	TREJO, SG	2002	18	51	MARTINEZ, VC	2006	7
17	ROMERO, AA	2003	6	52	MARTINEZ, DA	2007	4
18	ALVAREZ, RL	2003	9	53	ROMAN, GI	2007	15
19	GARCIA, MM	2003	2	54	HERNANDEZ, AR	2007	3
20	SARMIENTO, TJ	2003	38	55	HERNANDEZ, CE	2007	11
21	SANCHEZ, HJ	2003	7	56	GONZALEZ, AL	2007	9
22	MIRANDA, JS	2003	22	57	FABELA, CO	2007	7
23	JUAREZ, RF	2003	23	58	LOPEZ, YI	2007	17
24	QUIANE, RJ	2003	8	59	CATALAN, SE	2007	10
25	ARAGON, PE	2003	4	60	VILLEGAS, JV	2007	8
26	CUEVAS, RA	2003	11	61	RODRIGUEZ, RG	2007	7
27	DAVALOS, RF	2003	8	62	GOMEZ, BJ	2007	6
28	VARELA, GF	2003	5	63	HERNANDEZ, M	2007	14
29	PAREDES, AD	2004	6	64	GARRO, LB	2007	26
30	RAMOS, RR	2004	20	65	VERASTEGUI, BK	2007	8
31	COUDER, CC	2004	18	66	OLGUIN, AL	2007	8
32	FLORES, ME	2004	1	67	HORTA, MJ	2007	5
33	JIMENEZ, CE	2004	5	68	ORTEGA, GE	2007	9
34	RIVERA, BE	2004	11	69	MONTES DE OCA,	2007	7
35	JIMENEZ, HH	2004	12	70	RAMIREZ, AK	2007	15
				1017			

Fuente: Elaboración Propia con base en datos del CIC.2007

Sung y Jung (2008) clasifica al análisis de citaciones en seis áreas, los cuales se muestran en el siguiente apartado.

### 2.4.2.1 La naturaleza de las citas

Los libros que citan las tesis en general todas pertenecen al área de ciencias de la computación.

### 2.4.2.2 Autocitación

En el caso de estudio, no se puede analizar este indicador, sin embargo se dará un ejemplo con el uso de la herramienta HistCite versión 8.5.26 y con una colección de ejemplo. La colección consta de 500 registros, buscando autores con citas, localizándose al autor *Thelwall* a quien en uno de sus trabajos es citado ocho veces. Figura 2.4.4

2.4. 4 Autor con ocho citas

File Analyses View Tools Help					HistCite™			
Colección de Muestra-500					List Totals: LCS 21, GCS 269			
List of 10 Records for Author Thelwall M					List span: 2001 - 2004			
Records: 500, Authors: 984, Journals: 219, Cited References: 18460, Words: 1468								
Yearly output   Document Type   Language   Institution   Institution with Subdivision   Country								
#	Date / Author / Journal	LCS	GCS	LCR	CR			
<b>2001</b>								
1	392 Thelwall M Extracting macroscopic information from Web links JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY. 2001 NOV; 52 (13): 1157-1168	8	66	8	65			
<b>2002</b>								
2	398 Thelwall M Evidence for the existence of geographic trends in university Web site interlinking JOURNAL OF DOCUMENTATION. 2002; 58 (5): 563-574	4	34	6	32			
3	399 Thelwall M Research dissemination and invocation on the Web ONLINE INFORMATION REVIEW. 2002; 26 (6): 413-420	1	11	6	57			
4	424 Thelwall M Conceptualizing documentation on the Web: An evaluation of different heuristic-based models for counting links between university Web sites JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY. 2002 OCT; 53 (12): 995-1005	3	42	7	47			
<b>2003</b>								
5	431 Wilkinson D, Harries G, Thelwall M, Price L Motivations for academic web site interlinking: evidence for the Web as a novel source of information on informal scholarly communication JOURNAL OF INFORMATION SCIENCE. 2003; 29 (1): 49-56	0	16	7	49			
6	432 Tang R, Thelwall M US academic departmental Web-site interlinking in the United States disciplinary differences LIBRARY & INFORMATION SCIENCE RESEARCH. 2003; 25 (4): 437-458	1	11	20	77			
7	435 Vaughan L, Thelwall M Scholarly use of the Web: What are the key inducers of links to journal Web sites? JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY. 2003 JAN 1; 54 (1): 29-38	2	43	8	45			
8	447 Thelwall M, Harries G The connection between the research of a university and counts of links to its web pages: An investigation based upon a classification of the relationships of pages to the research of the host university JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY. 2003 MAY; 54 (7): 594-602	1	17	8	56			
<b>2004</b>								
9	460 Harries G, Wilkinson D, Price L, Fairclough R, Thelwall M Hyperlinks as a data source for science mapping JOURNAL OF INFORMATION SCIENCE. 2004; 30 (5): 436-447	0	11	11	45			
10	463 Thelwall M, Harries G Do the Web sites of higher rated scholars have significantly more online impact? JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY. 2004 JAN 15; 55 (2): 149-159	1	18	17	101			

Fuente: Elaboración Propia con base en datos del CIC, utilizando la herramienta HistCite.2007

Con la misma herramienta se revisó de quienes provienen las ocho citas, obteniéndose la Figura 2.4.5 Donde se puede apreciar que las ocho publicaciones hacen referencia al trabajo mostrado en la Figura 2.4.4 que pertenecen al mismo autor.

2.4.5 Publicaciones que referencias a la publicación 392

File Analyses View Tools Help HistCite™  
List Totals: LCS 13, GCS 192  
List span: 2002 - 2004

**Colección de Muestra-500**  
List of 8 Records citing 392 Thelwall M  
Extracting macroscopic information from Web links  
JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY. 2001  
NOV; 52 (13): 1157-1168

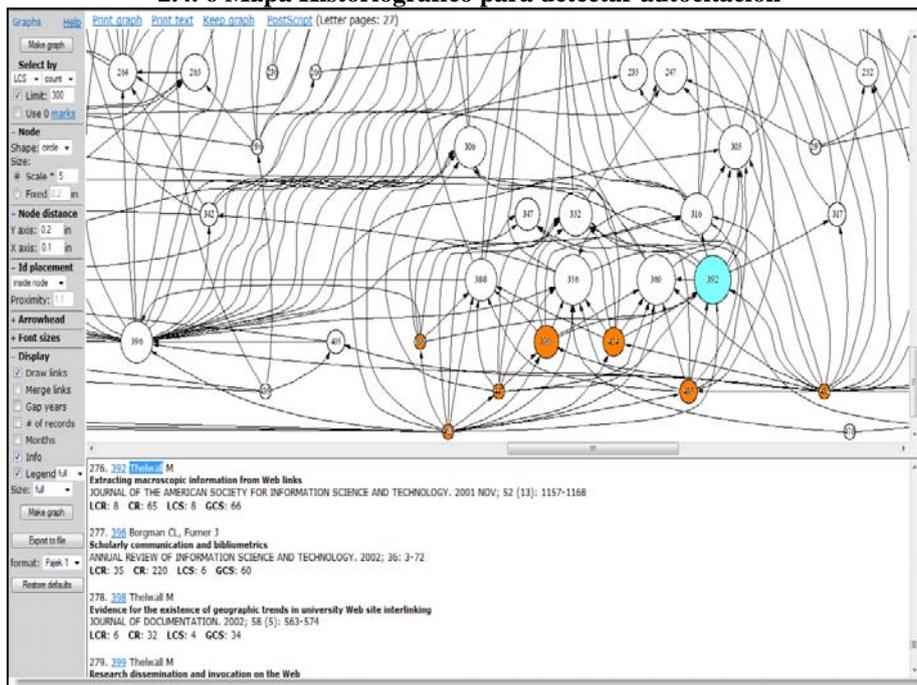
Records: 500, Authors: 984, Journals: 219, Cited References: 18460, Words: 1468  
Yearly output | Document Type | Language | Institution | Institution with Subdivision | Country

#	Date / Author / Journal	LCS	GCS	LCR	CR
<b>2002</b>					
1	398 Thelwall M Evidence for the existence of geographic trends in university Web site interlinking JOURNAL OF DOCUMENTATION. 2002; 58 (5): 563-574	4	34	6	32
2	399 Thelwall M Research dissemination and invocation on the Web ONLINE INFORMATION REVIEW. 2002; 26 (6): 413-420	1	11	6	57
3	424 Thelwall M Conceptualizing documentation on the Web: an evaluation of different heuristic-based models for counting links between university Web sites JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY. 2002 OCT; 53 (12): 995-1005	3	42	7	47
<b>2003</b>					
4	431 Wilkinson D, Harries G, Thelwall M, Price L Motivations for academic web site interlinking: evidence for the Web as a novel source of information on informal scholarly communication JOURNAL OF INFORMATION SCIENCE. 2003; 29 (1): 49-56	0	16	7	49
5	432 Tang R, Thelwall M US academic departmental Web-site interlinking in the United States disciplinary differences LIBRARY & INFORMATION SCIENCE RESEARCH. 2003; 25 (4): 437-458	1	11	20	77
6	435 Vaughan L, Thelwall M Scholarly use of the web: what are the key inducers of links to journal Web sites? JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY. 2003 JAN 1; 54 (1): 29-38	2	43	8	45
7	447 Thelwall M, Harries G The connection between the research of a university and counts of links to its web pages: An investigation based upon a classification of the relationships of pages to the research of the host university JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY. 2003 MAY; 54 (7): 594-602	1	17	8	56
<b>2004</b>					
8	463 Thelwall M, Harries G Do the web sites of higher rated scholars have significantly more online impact? JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY. 2004 JAN 15; 55 (2): 149-159	1	18	17	101

Fuente: Elaboración Propia con base en datos del CIC, utilizando la herramienta HistCite.2007

Esto mismo se puede observar con un mapa historiográfico, donde se aprecian con esferas que representan las publicaciones 398, 399, 424, 431, 432, 435, 447 y 463, que pertenecen al mismo autor *Thelwall* y que además citan a la publicación 392 del mismo autor. Figura 2.4.6

2.4.6 Mapa Historiográfico para detectar autocitación



Fuente: Elaboración Propia con base en datos de HistCite.2007

### 2.4.2.3 Libros altamente citados

En la Figura 2.4.7 sobresale el libro más citado, siendo este el Manual de XML.

**2.4.7 Listado de Libros más citados**

#	Date / Author / Journal	LCS	GCS	LCR	CR
1	642 GOLDFARB CF, PRESCOD PL <b>MANUAL DE XML</b> PRENTICE HALL. 1999;	5		0	0
2	200 KOHONEN T <b>SELF-ORGANIZATION AND ASSOCIATIVE MEMORY</b> SPRINGER-VERLAG. 1989;	4		0	0
3	301 HASSOUN MH <b>ASSOCIATIVE NEURAL MEMORIES</b> OXFORD UNIVERSITY PRESS. 1993;	4		0	0
4	355 GAMMA E, HELM R, JOHNSON R, VLISSIDES J <b>DESIGN PATTERNS ELEMENTS OF REUSABLE OBJECT-ORIENTED SOFTWARE</b> ADDISON WESLEY PUBLISHING COMPANY. 1995;	4		0	0
5	356 DATE CJ <b>AN INTRODUCTION TO DATABASE SYSTEMS</b> ADDISON WESLEY PUBLISHING COMPANY. 1995;	4		0	0
6	363 ALLEN JF <b>NATURAL LANGUAGE UNDERSTANDING</b> BENJAMIN CUMMINGS PUBLISHING COMPANY INC. 1995;	4		0	0
7	547 PRESSMAN R	4		0	0

Fuente: Elaboración Propia con base en datos del CIC, utilizando la herramienta HistCite.2007

### 2.4.2.4 Análisis sobre campos especiales

En las siguientes tres graficas, se muestra análisis por tres diferentes campos. La Figura 2.4.8 Análisis por tipo Recursos de Información, La Figura 2.4.9 palabras clave en títulos de libros y tesis, la Figura 2.4.10 por idioma de citaciones, La Figura 2.4.11 por país y la Figura 2.4.12 por años de publicación.

Figura 2.4. 1 Recursos de Información.

#	Document Type	Recs	TLCS	TGCS
1	BOOK	908	1017	
2	THESES	70	0	

Fuente: Elaboración Propia con base en datos del CIC, utilizando la herramienta HistCite.2007

En la Figura 2.4.9 se muestra una estadística de las palabras que aparecen en los títulos de los libros y tesis citadas por los alumnos del centro, ordenadas de mayor a menor frecuencia, cabe mencionar que dichas palabras no se encuentran homogenizadas. Por ejemplo Sistemas, Systems, Sistema, esto significa que pueden haber palabras que se escriben de forma diferente y significan lo mismo, expresiones o palabras con una misma raíz que se puedan agrupar en una sola.

En esta estadística se observa que la palabra mas citada es **java**, este valor indica la frecuencia del lenguaje de programación y es de esperarse ya que la mayoría de las tesis son desarrollo de aplicaciones. En segundo lugar figura la palabra **sistema**, en la mayoría de las tesis los alumnos hacen entrega de un sistema o trabajan en uno o varios sistemas operativos, ya que en las tesis hacen uso de un lenguaje de **programación**, que seguramente entregan un **software** como producto final.

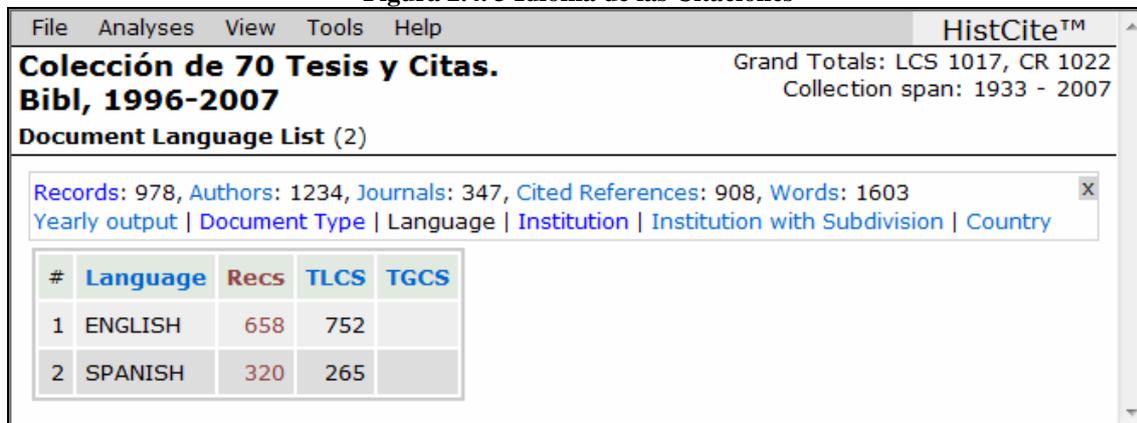
**Figura 2.4. 2 Palabras clave en títulos de libros y tesis citados**

#	Word	Recs	TLCS	TGCS
1	JAVA	73	88	
2	SYSTEMS	56	67	
3	PARA	55	22	
4	SOFTWARE	49	56	
5	WEB	45	46	
6	XML	42	59	
7	PROGRAMMING	38	47	
8	SISTEMAS	37	33	
9	THEORY	37	39	
10	DATOS	36	25	
11	CON	34	29	
12	INTRODUCTION	30	36	
13	ANALISIS	27	24	
14	APPLICATIONS	27	28	
15	OBJETOS	27	25	
16	COMPUTER	26	26	
17	DATA	25	26	

Fuente: Elaboración Propia con base en datos del CIC, utilizando la herramienta HistCite.2007

Se observa un dominio en el idioma inglés en cuanto a libros que citan los alumnos, sin embargo el centro no cuenta con ninguna tesis traducida a un idioma diferente al español durante el periodo en análisis. Figura 2.4.10.

Figura 2.4. 3 Idioma de las Citaciones



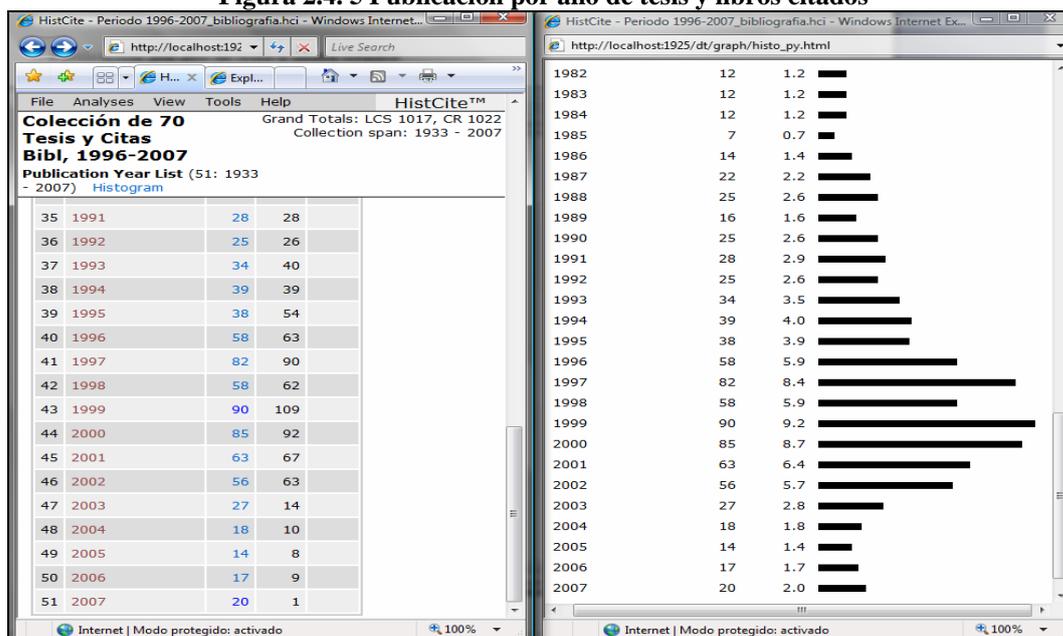
Fuente: Elaboración Propia con base en datos del CIC, utilizando la herramienta HistCite.2007

Figura 2.4. 4 Por país



Fuente: Elaboración Propia con base en datos del CIC, utilizando la herramienta HistCite.2007

Figura 2.4. 5 Publicación por año de tesis y libros citados



Fuente: Elaboración Propia con base en datos del CIC, utilizando la herramienta HistCite.2007

### 2.4.3.5 Usando aspectos

Se relaciona con aspectos sociales y culturales de como se llevan a cabo las citas. Este índice queda fuera de este estudio.

### 2.4.3.6 Análisis de Cocitación

Para el análisis de cocitación, se busco un autor que sea citado más de una vez, seleccionando al autor *Goldfarb*. Figura 2.4.13, el cual es citado tres veces.

Figura 2.4. 6 Autor *Goldfarb* con tres citas

#	Date / Author / Journal	LCS	GCS	LCR	CR
1	114 Goldfarb CF, Prescod P <b>Manual de XML</b> PRENTICE HALL. 1999;	3		0	0
2	52 Signore R <b>The ODBC Solution</b> MCGRAW HILL. 1995;	2		0	0
3	156 Young M <b>XML STEP BY STEP</b> MICROSOFT PRESS. 2000;	2		0	0
4	1 Morik K, Wrobel S, Kietz JU, Emde W <b>Knowledge Acquisition and Machine Learning:Theory, Methods and Applications</b> ACADEMIC PRESS. 1933;	1		0	0
5	2 Minsky M, Papert S <b>Perceptrons</b> MIT PRESS. 1969;	1		0	0
6	3 Box GEP, Jenkins GM <b>Time Series Analysis, Forecasting and Control</b> WILEY. 1970;	1		0	0
7	4 Watts DJ <b>Six Degrees: The science of a connected age</b> NORTON & COMPANY INC. 1971;	1		0	0
8	5 Codd EF <b>Further Normalization of the DataBase Relational Model, en Data Base Systems Courant Computer Science Symposio Series</b> PRENTICE HALL. 1972;	1		0	0

Fuente: Elaboración Propia con base en datos del CIC, utilizando la herramienta HistCite.2007

Posteriormente se analizaron de los tres temas de tesis ver (Figura 2.4.13) sus correspondientes bibliografías que citan, y se pudo ver que tienen una co-citación en varios autores (*Goldfarb*, y *Date*). Figura 2.4.14.

Figura 2.4. 7 Tres temas de tesis que citan a Goldfarb

#	Date / Author / Journal	LCS	GCS	LCR	CR
<b>2002</b>					
1	259 Trejo SG Modulo XML para acceder al sistema administrador de bases de datos SQLMx a traves de Internet CIC-IPN. 2002 JUN 14;	0		18	18
<b>2004</b>					
2	307 Rivera BE Conversion de archivos con registros arborescentes de texto (RAT) a base de datos relacional CIC-IPN. 2004 OCT 14;	0		11	11
<b>2005</b>					
3	321 Gonzalez EL Interfaz para la publicacion de las operaciones de consulta en un sistema administrador de bases de datos (SABD) como servicio web CIC-IPN. 2005 MAY 04;	0		49	49

Fuente: Elaboración Propia con base en datos del CIC, utilizando la herramienta HistCite.2007

Figura 2.4. 8 Co-citación en dos autores por parte de tres temas de tesis

Record ID	Author(s)	Title	Source	Date	Document Type	Language	LCR	CR	LCS	GCS	Cited References (Highlighted)
259	Trejo SG	Modulo XML para acceder al sistema administrador de bases de datos SQLMx a traves de Internet	CIC-IPN	2002 JUN 14	Theses	Spanish	18	18	0		GOLDFARB CF, 1999, PRENTICE HALL,,, DATE CJ, 1999, ADDISON WESLEY,,,
307	Rivera BE	Conversion de archivos con registros arborescentes (RAT) a base de datos relacional	CIC-IPN	2004 OCT 14	Theses	Spanish	11	11	0		GOLDFARB CF, 1999, PRENTICE HALL,,, DATE CJ, 1999, ADDISON WESLEY,,,
321	Gonzalez EL	Interfaz para la publicacion de las operaciones de consulta en un sistema administrador de bases de datos (SABD) como servicio web	CIC-IPN	2005 MAY 04	Theses	Spanish	49	49	0		GOLDFARB CF, 1999, PRENTICE HALL,,, DATE CJ, 1999, ADDISON WESLEY,,,

Fuente: Elaboración Propia con base en datos del CIC, utilizando la herramienta HistCite.2007

### 2.4.4 Índice de Impacto

El cálculo del índice de impacto se realizó con el autor *Booch G*, que tiene doce publicaciones y quince citas. Figura 2.4.16.

Figura 2.4. 9 Autor *Booch G* con doce publicaciones y quince citas

The screenshot shows the HistCite 2007 interface. At the top, it displays 'Colección de 70 Tesis y Citas Bib, 1996-2007' and 'All-Author List (1232)'. Below this, there are statistics: 'Records: 977, Authors: 1232, Journals: 347, Cited References: 907, Words: 1603'. A table lists authors with columns for '#', 'Author', 'Recs', 'TLCS', and 'TGCS'. The first row, for 'BOOCH G', is circled in red, showing 12 publications and 15 citations.

#	Author	Recs	TLCS	TGCS
1	BOOCH G	12	15	
2	PRESSMAN RS	11	17	
3	RUMBAUGH J	9	11	
4	CHOMSKY N	6	7	
5	GONZALEZ RC	6	7	
6	JACOBSON I	6	8	
7	ULLMAN JD	6	6	
8	AHO AV	5	5	
9	DEITEL HM	5	8	
10	SILBERSCHATZ A	5	5	
11	DATE CJ	4	10	
12	DEITEL PJ	4	6	

Fuente: Elaboración Propia con base en datos del CIC, utilizando la herramienta HistCite.2007

El cálculo del índice de impacto del autor *Booch G*, es el siguiente:

$$I_i = \text{Total de citas} / \text{Total de publicaciones} = 15/12 = 1.25$$

# Capítulo 3

## Aplicación del método de las palabras asociadas

En el presente capítulo se describe la manera en que se llevó a cabo el análisis de las palabras asociadas, con la finalidad de describir las relaciones entre los tópicos o conceptos de las Ciencias de la Computación que se toman como temas en los desarrollos de tesis. Se consideran las palabras tanto del resumen como del título de 90 tesis, divididas en tres periodos de tiempo, siendo estos 1996-2001; 2002-2004; 2005-2007.

Para esto en el punto 3.1 detalla como se llegó a formar la metodología para este análisis; en el punto 3.2 se describe como se utiliza en cierto nivel de detalle el funcionamiento y los resultados parciales al utilizar el software Redes 2005, software que nos permite hacer el análisis utilizando el método de las palabras asociadas y en el punto 3.3 se describe el análisis en los tópicos de las ciencias de la computación.

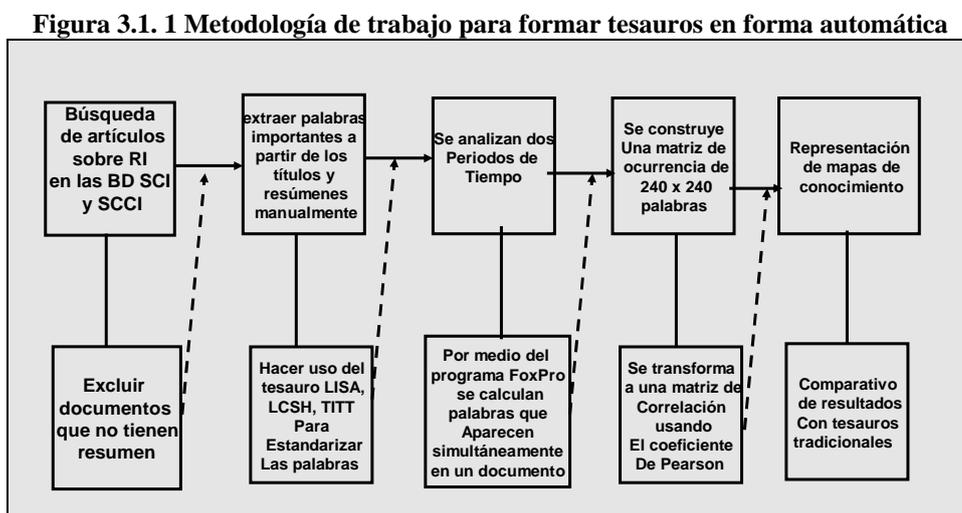
### 3.1 Selección de Metodología de Análisis

La labor de buscar patrones y tendencias en los tópicos de las Ciencias de la Computación en los desarrollos de temas de tesis de maestría, se efectuó con las palabras de los títulos y resúmenes, lo anterior se definió después de revisar una serie de casos de estudio en donde aplican el método de las palabras asociadas Wyong (2007), Ding (2000), Chowdhury (2000) y Foo (2000) con objetivos similares y sus correspondientes metodologías, como las que se describen brevemente en los párrafos siguientes.

#### 3.1.1 Organizando las palabras claves en un ambiente de búsqueda en la web: Una metodología basada en el análisis de palabras asociadas

En este artículo Ding et al., (2000) lleva a cabo el análisis de palabras asociadas, donde se puede generar en forma automática un tesoro aprovechando la característica de agrupamientos naturales. El análisis se desarrolló en artículos que se obtuvieron de las bases de datos Science Citation Index (SCI) y Social Science Citation Index (SSCI) cubriendo un periodo de 1987-1997, las palabras para los agrupamientos se extrajeron

manualmente de títulos y resúmenes en el campo de recuperación de la información (RI). En los resultados de otro caso de estudio aplicando la misma metodología Ding, Chowdhury y Foo (2001) mostraron que el campo de la RI tiene temas de investigación establecidas y que cambia rápidamente hacia nuevos temas. La metodología a grandes rasgos. Figura 3.1.1:

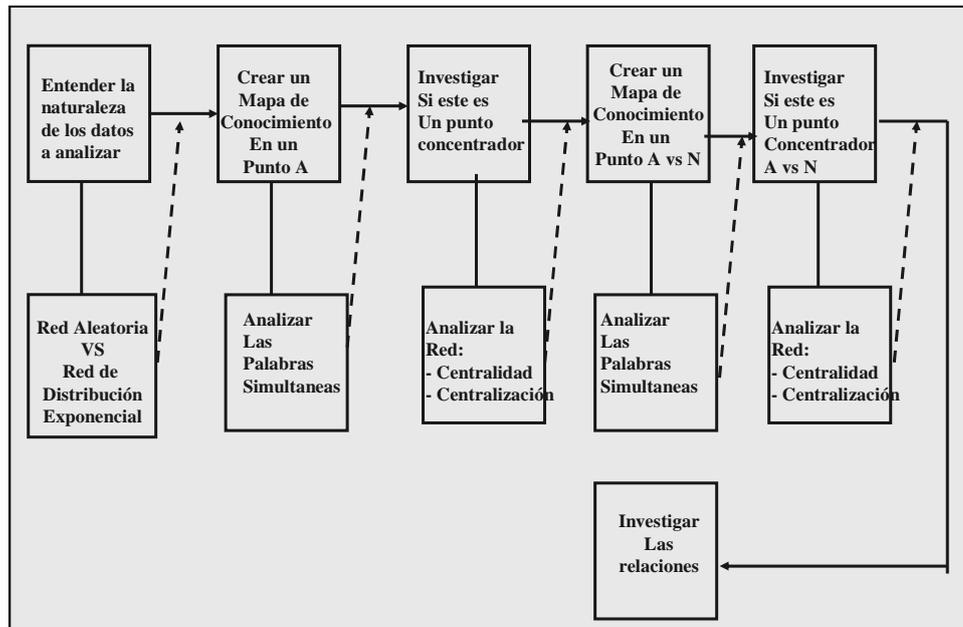


Fuente: Ding, Chowdhury y Foo.2001

### 3.1.2 Como identificar campos de investigación emergentes usando cienciometría: Un ejemplo en el campo de la seguridad de la información.

En este artículo Hyoun (2007) realiza un análisis del método de las “palabras asociadas”, el cual fue empleado para localizar patrones y tendencias en el campo de la Seguridad de la Información para medir la fuerza de la asociación de términos representativos de publicaciones relevantes u otros textos producidos en dicho campo. Se trabajan con tesis de doctorado las cuales se obtuvieron de la base de datos de Science Citation Index (SCI) y las palabras claves se extrajeron a partir del criterio de expertos en el área, posteriormente fueron estandarizadas. Para darle un seguimiento a los cambios dinámicos en los campos de la Seguridad de la Información, se presentan dos mapas tecnológicos. Los resultados muestran que el campo de la Seguridad de la Información tiene algunos temas de investigación ya establecidos y fueron resumidas en seguridad de la información, computación y criptosistemas. Todo a través del debate con profesores en campos relacionados, criptosistema resultó tener una gran posibilidad de convertirse en el tema mas popular en los próximos años. La metodología de trabajo se muestra en la Figura 3.1.2.

**Figura 3.1. 2 Metodología de trabajo para localizar tendencias en un campo de conocimiento**



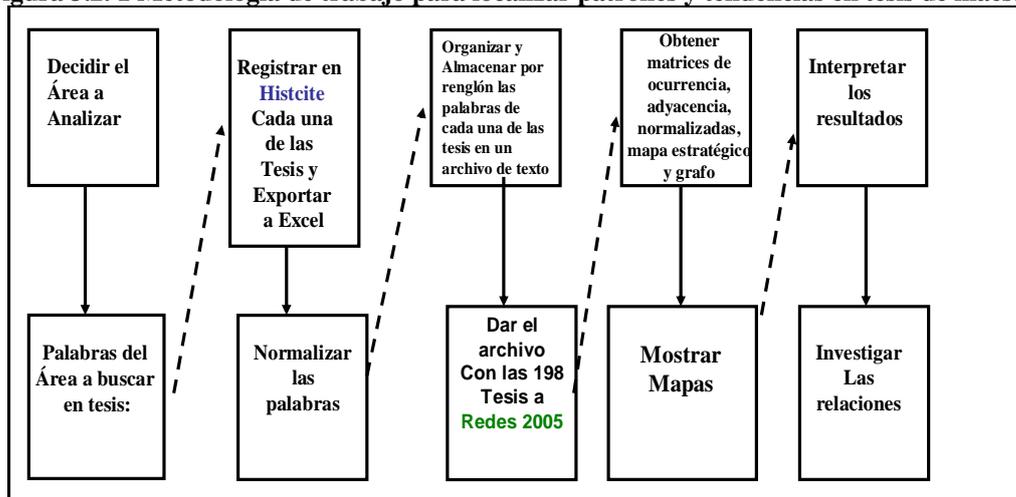
Fuente: Elaboración Propia a partir de Hyoung, 2007

Como se ha visto en los artículos revisados en donde se persiguen objetivos similares, se hace uso del método de las palabras asociadas, sin embargo en ellas no describen en detalle los pasos que se realizan. En el siguiente apartado se describe la metodología utilizada en este trabajo de tesis.

### 3.2 Descripción de la Metodología Utilizada.

Para localizar los patrones y tendencias en los desarrollos de temas de tesis, la estructura de la metodología desarrollada se puede ver en la Figura 3.2.1, resultado del análisis de las metodologías anteriores. El detalle de algunos de los puntos, como el cálculo de las matrices y los índices de equivalencia utilizado se obtuvo de (<http://www.ugr.es/~rruizb/cognosfera/>)

Figura 3.2. 1 Metodología de trabajo para localizar patrones y tendencias en tesis de maestría



Fuente: Elaboración Propia. 2008

A continuación se describen los pasos de la metodología utilizada:

### 3.2.1 Decidir el texto y el área a analizar.

Dentro de la estructura de los documentos de las tesis en estudio se analiza el texto del título y resumen, se toma el título debido a que hace mención al nombre del trabajo desarrollado y posiblemente algún tópico de las ciencias de la computación y en el resumen los autores generalmente hacen mención a tópicos relevantes.

Para la búsqueda de patrones y tendencias se tomaron en cuenta todos los tópicos que abarcan los documentos de temas de tesis en estudio, pues en términos generales el número de documentos a tratar son relativamente pocos y si se eligiera una sola área de las ciencias de la computación estos números de documentos a analizar se reduciría más.

Se analizaron los documentos de tesis y se aprovecho una característica del método de las palabras asociadas que es, que sin indicar un área en específico los agrupamientos que se obtengan sea en forma natural y al ser visualizados en el diagrama estratégico y el grafo correspondiente indique una descripción visual sobre los tópicos que se desarrollaron en los temas de tesis.

### 3.2.2 Palabras del área a buscar en tesis.

En este paso generalmente las palabras claves se normalizan por medio de alguna herramienta informática, tal que las palabras claves puede tener consistencia (singular y plural), unicidad (sinónimos) y claridad (homónimos) también por los términos de los tesauros de tecnologías de la información.

En esta tesis la definición de estas se realizó con ayuda de los investigadores del Centro El Dr. Adolfo Guzmán Arenas y el Dr. Gilberto Martínez Luna, tomándose como referencia la currícula del la ACM (Association for Computing Machinery) of Computer Science y su experiencia.

### 3.2.3 Registrar en HistCite versión 8.5.26 cada una de las tesis y exportar a Excel

Utilizando la capacidad de esta herramienta para separar las palabras del título y del resumen; se registraron cada una de las tesis y posteriormente se exportaron a un archivo para ser utilizado por Excel. En este paso la herramienta HistCite versión 8.5.26 eliminaba las palabras de uno y dos caracteres.

Cabe aclarar, que para poder manejar un contexto denotado por varias palabras, por ejemplo *bases de datos* (son tres palabras), estas palabras se unieron por medio de un - (*guión*), con lo cual HistCite versión 8.5.26 las trataba como una sola palabra, es decir la palabra *bases-de-datos*.

### 3.2.4 Normalizar las palabras

En este paso se normalizaron las palabras de la siguiente forma:

- 1) Se eliminaban las palabras sin sentido en el campo de las Ciencias de la Computación. Por ejemplo: acerca, acrecentar, aunado, ayuda, entre otros.
- 2) Con ayuda de un experto en el área de las Ciencias de la Computación se analizaron cada una de las palabras en Excel y se identificaba la palabra que mejor describía el tema

de la tesis, si era la misma se dejaba en caso contrario se remplazaba, es decir no perder el contexto del tema de tesis.

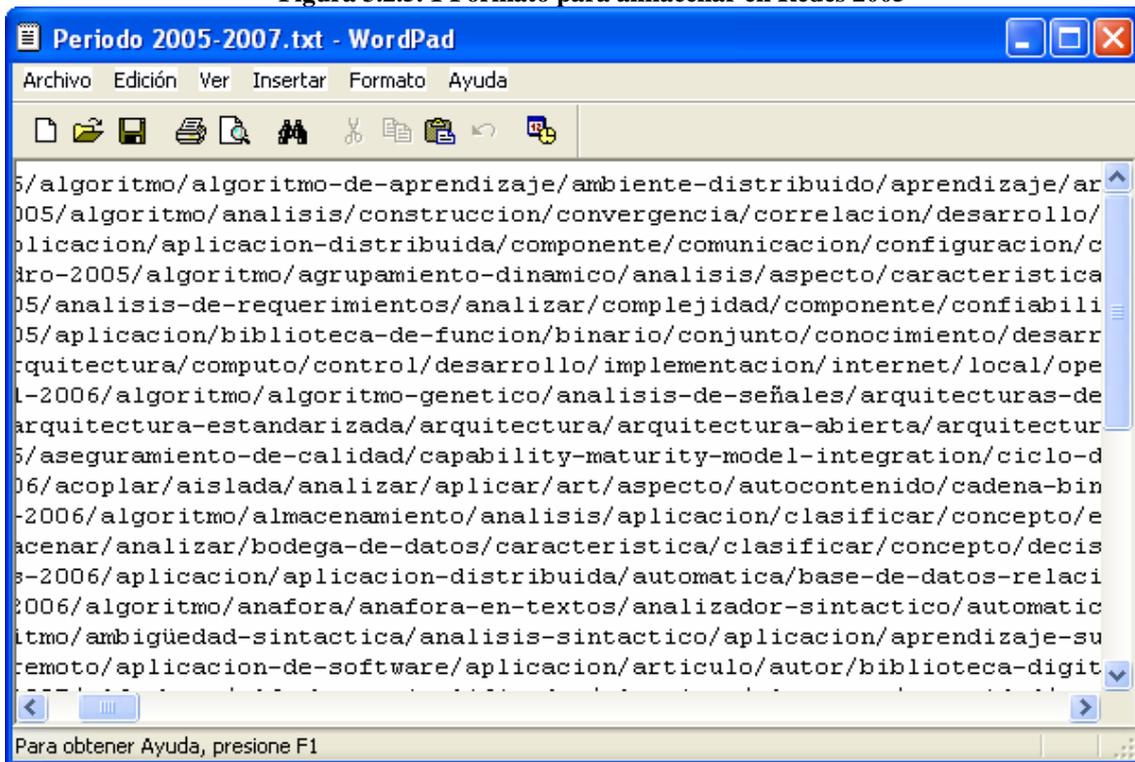
Ejemplo:

ejecución, ejecutar, ejecutado → ejecutar  
 bases de datos, base de datos → base de datos  
 actualizacion, actualiza, actualizar → actualizar

### 3.2.5 Organizar y almacenar por renglón cada una de las tesis en un archivo

La herramienta Redes 2005, requiere de un formato para poder procesar las palabras de cada una de las tesis, la Figura 3.2.5.1 muestra el formato que se requiere por cada una.

Figura 3.2.5. 1 Formato para almacenar en Redes 2005

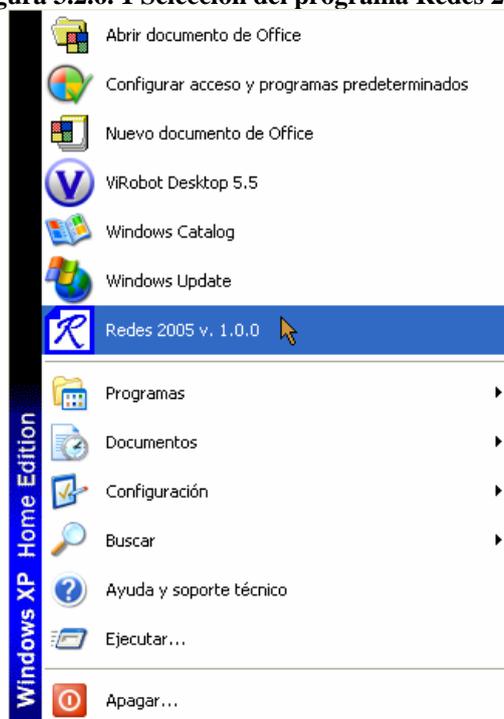


Fuente: Elaboración Propia con base en datos del CIC. 2008

### 3.2.6 Dar el archivo con las 90 tesis a Redes 2005

Los pasos para el método de las palabras asociadas se llevo a cabo con la ejecución del programa Redes 2005, en las siguientes figuras se ilustra la creación de cada una de las sesiones de análisis: la selección del archivo con las palabras a analizar y la captura de los parámetros necesarios para seleccionar las palabras que formaran los grupos y las que serán las representantes de cada uno de los grupos, así como las matrices necesarias para los cálculos de los índices de equivalencia y finalmente llegar a la visualización de los grupos. Las figuras siguientes representan este proceso para el periodo 1996-2007.

**Figura 3.2.6. 1 Selección del programa Redes 2005.**



**Fuente: Ventana de Inicio.2008**

Para crear un nuevo proyecto, se selecciona “nuevo” del menú principal de Redes 2005

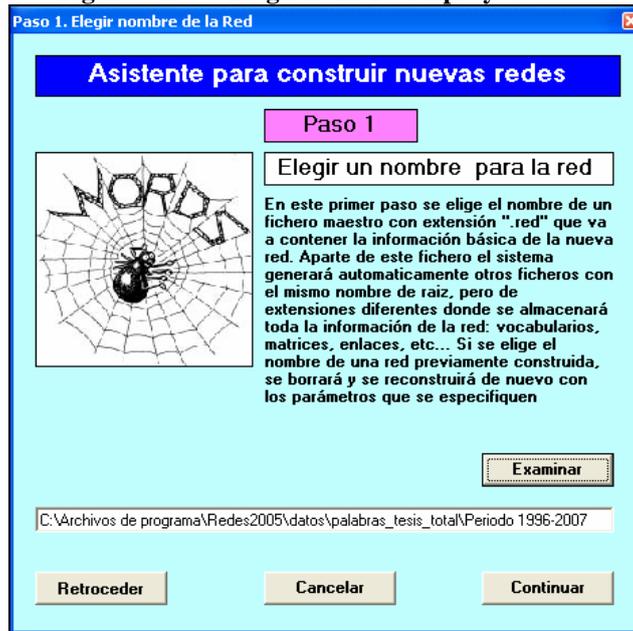
**Figura 3.2.6. 2 Selección de nuevo proyecto.**



**Fuente: Redes 2005 v.1.0.0.2008**

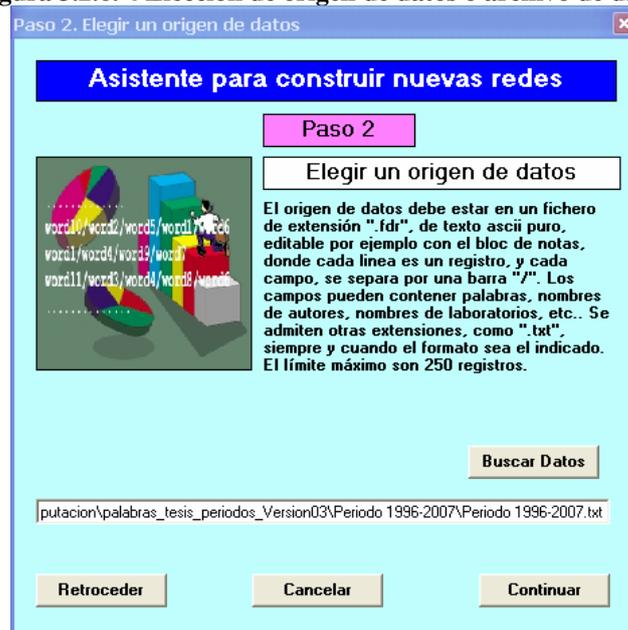
El siguiente paso es elegir el nombre de la red *Periodo 1996-2007*, posteriormente el origen de los datos y por último los parámetros de construcción de las redes.

**Figura 3.2.6. 3 Elegir nombre del proyecto.**



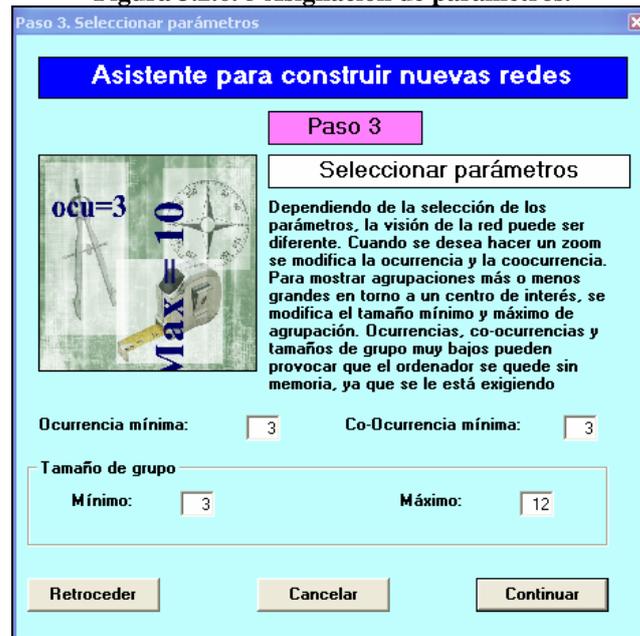
Fuente: Redes 2005 v.1.0.0.2008

**Figura 3.2.6. 4 Elección de origen de datos o archivo de datos.**



Fuente: Redes 2005 v.1.0.0.2008

Figura 3.2.6. 5 Asignación de parámetros.



Fuente: Redes 2005 v.1.0.0.2008

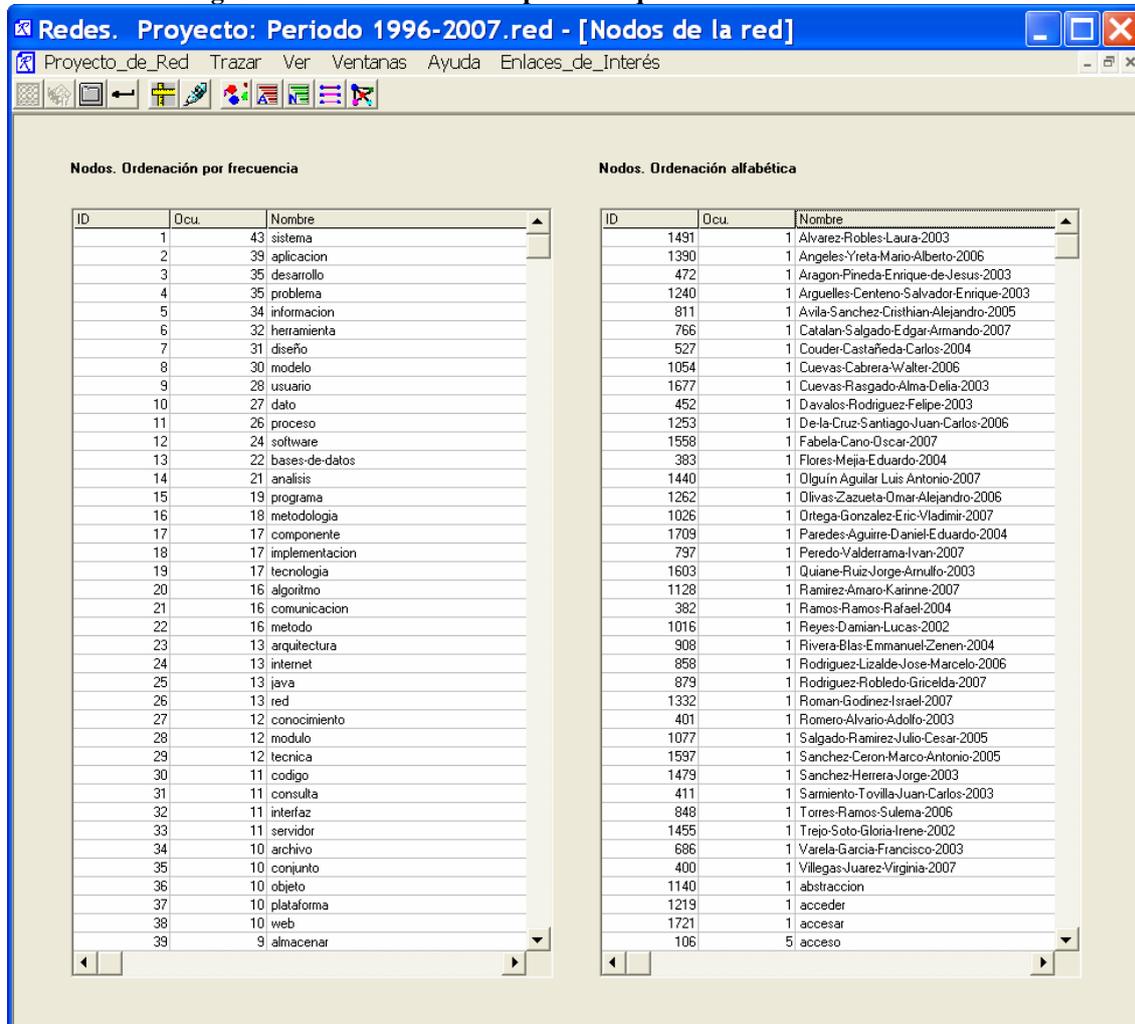
Figura 3.2.6. 6 Número de palabras en análisis y palabras que superan el umbral.



Fuente: Redes 2005 v.1.0.0.2008

En la Figura 3.2.6.7 se muestran las palabras por orden de frecuencia y alfabéticamente. Se puede observar que las palabras con más frecuencia son: *sistema*, *aplicación*, *desarrollo* y *problema*, solo por mencionar algunas.

**Figura 3.2.6. 7 Listado de las palabras que se localizan en el análisis.**



Fuente: Elaboración Propia con base de datos del CIC usando el software Redes 2005 v.1.0.0.2008

### 3.2.7 Obtener Matrices de adyacencia, normalizadas, mapa estratégico y grafo.

La matriz de adyacencia muestra los valores de co-ocurrencias de los nodos en los documentos.

**Figura 3.2.7. 1 Matriz de Adyacencia. Co-ocurrencia de Palabras**

	1 sistema	2 aplicador	3 desarrollo	4 problema	5 informac	6 herramien	7 diseño	8 modelo	9 usuario	10 dato	11 proceso	12 software	13 bases-di	14 analisis	15 program
1 sistema	43	19	19	14	20	17	20	13	15	16	13	15	15	10	10
2 aplicador		39	18	15	20	14	16	14	15	13	12	11	13	11	8
3 desarrollo			35	16	15	14	13	11	13	14	14	12	11	6	6
4 problema				35	12	13	14	12	8	8	14	7	6	10	6
5 informac					34	17	11	12	16	13	11	6	12	5	6
6 herramien						32	13	13	11	10	10	11	8	9	7
7 diseño							31	13	13	9	11	9	7	9	9
8 modelo								30	10	6	7	12	6	8	5
9 usuario									28	12	8	9	7	5	10
10 dato										27	9	8	7	4	9
11 proceso											26	5	6	7	7
12 software												24	5	5	7
13 bases-di													22	4	2
14 analisis														21	4
15 program															19
16 metodol															

Fuente: Elaboración Propia con base de datos del CIC usando el software Redes 2005 v.1.0.0.2008

En la Figura 3.2.7.1 se muestra la matriz de adyacencia, esta es válida para trazar las redes si esta normalizada. El método de las palabras asociadas emplea el índice de equivalencia para realizar esta normalización. Se define como el cociente entre el cuadrado de la co-ocurrencia entre los nodos i y j y el producto de las ocurrencias de estos por separado. (<http://www.ugr.es/~rruizb/cognosfera/>). Figura 3.2.7.2

**Figura 3.2.7. 2 Matriz Normalizada con el índice de equivalencia.**

	1 sistema	2 aplicador	3 desarrollo	4 problema	5 informac	6 herramien	7 diseño	8 modelo	9 usuario	10 dato	11 proceso	12 software	13 bases-di	14 analisis	15 program
1 sistema	1	0.215	0.24	0.13	0.274	0.21	0.3	0.131	0.187	0.22	0.151	0.218	0.238	0.111	0.122
2 aplicador		1	0.237	0.165	0.302	0.157	0.212	0.168	0.206	0.16	0.142	0.129	0.197	0.148	0.086
3 desarrollo			1	0.209	0.189	0.175	0.156	0.115	0.172	0.207	0.215	0.171	0.157	0.049	0.054
4 problema				1	0.121	0.151	0.181	0.137	0.065	0.068	0.215	0.058	0.047	0.136	0.054
5 informac					1	0.266	0.115	0.141	0.269	0.184	0.137	0.044	0.193	0.035	0.056
6 herramien						1	0.17	0.176	0.135	0.116	0.12	0.158	0.091	0.121	0.081
7 diseño							1	0.182	0.195	0.097	0.15	0.109	0.072	0.124	0.138
8 modelo								1	0.119	0.044	0.063	0.2	0.055	0.102	0.044
9 usuario									1	0.19	0.088	0.121	0.08	0.043	0.188
10 dato										1	0.115	0.099	0.082	0.028	0.158
11 proceso											1	0.04	0.063	0.09	0.099
12 software												1	0.047	0.05	0.107
13 bases-di													1	0.035	0
14 analisis														1	0.04
15 program															1

Fuente: Elaboración Propia con base de datos del CIC usando el software Redes 2005 v.1.0.0.2008

En la Figura 3.2.7.3 se muestra la lista de enlaces ordenados por el índice de equivalencia del periodo 1996-2007.

Figura 3.2.7. 3 Lista de enlaces ordenados por índice de equivalencia.

Enlaces ordenados por índice de equivalencia										
Rango	Selec.	ID1	Ocu	Nombre	ID2	Ocu	Nombre	Coocu	e	
1	*	104	5	sistema-administrador-de-bases-c	183	3	sentencia	3	0.600	
2		52	8	concepto	111	4	clasificar	4	0.500	
3	*	75	6	lenguaje-de-programacion	193	3	url	3	0.500	
4	*	85	6	sql	183	3	sentencia	3	0.500	
5	*	98	5	ingenieria-de-software	116	4	desarrollo-de-software	3	0.450	
6	*	63	7	ejecutar	163	3	actualizar	3	0.429	
7	*	33	11	servidor	50	8	cliente	6	0.409	
8	*	68	7	relacion	76	6	nodo	4	0.391	
9	*	49	8	calculo	141	3	adquisicion-de-datos	3	0.375	
10	*	49	8	calculo	191	3	tres-dimensiones	3	0.375	
11	*	81	6	prototipo	116	4	desarrollo-de-software	3	0.375	
12	*	18	17	implementacion	23	13	arquitectura	9	0.367	
13	*	98	5	ingenieria-de-software	100	5	marco	3	0.360	
14	*	42	9	construir	150	3	agente	3	0.333	
15	*	25	13	java	75	6	lenguaje-de-programacion	5	0.321	
16	*	63	7	ejecutar	112	4	comando	3	0.321	
17	*	2	39	aplicacion	5	34	informacion	20	0.302	
18	*	1	43	sistema	7	31	diseño	20	0.300	
19	*	34	10	archivo	141	3	adquisicion-de-datos	3	0.300	
20	*	38	10	web	193	3	url	3	0.300	
21	*	70	6	ambiente	93	5	entorno	3	0.300	
22	*	77	6	operacion	104	5	sistema-administrador-de-bases-c	3	0.300	
23	*	81	6	prototipo	98	5	ingenieria-de-software	3	0.300	
24	*	82	6	proyecto	98	5	ingenieria-de-software	3	0.300	
25	*	82	6	proyecto	100	5	marco	3	0.300	
26	*	85	6	sql	104	5	sistema-administrador-de-bases-c	3	0.300	
27	*	43	9	control	70	6	ambiente	4	0.296	
28	*	50	8	cliente	60	7	cliente-servidor	4	0.286	
29	*	57	8	lenguaje	60	7	cliente-servidor	4	0.286	
30	*	3	35	desarrollo	19	17	tecnologia	13	0.284	
31	*	49	8	calculo	113	4	computadora-personal	3	0.281	
32	*	54	8	funcion	112	4	comando	3	0.281	
33	*	58	8	parametro	114	4	configuracion	3	0.281	
34	*	58	8	parametro	120	4	grafico	3	0.281	
35	*	1	43	sistema	5	34	informacion	20	0.274	

Fuente: Elaboración Propia con base de datos del CIC usando el software Redes 2005 v.1.0.0.2008

En la Figura 3.2.7.4, se muestra la ficha informativa de los agrupamientos, en este caso del agrupamiento “sistema”, se muestra un conjunto de valores cuantitativos como enlaces internos y externos, así como el grado de centralidad y densidad.



### **3.2.9 Interpretar los Resultados.**

Para el periodo 1996 – 2007 los temas que se encuentran en el **cuadrante 1**, *sistema*, *Internet* y *servidor* son temas centrales en la red general y que se encuentran relacionados con otros temas. Estos temas constituyen de alguna forma el centro del campo, sin embargo para los temas que se muestran en el **cuadrante 2** deben de prestarse atención, debido a que son susceptibles de convertirse en centrales y desarrollados y consiguientemente desplazarse hacia el primer cuadrante. El **cuadrante 3** se muestra el tema *modelo* probablemente en una fase anterior era central y se ha visto marginado, en el ultimo **cuadrante 4** se muestran los temas *red*, *estructura*, *algoritmo* y *objeto* los cuales son temas que van emergiendo o que tienden a desaparecer, para ello es necesario estudiar cortos periodos de tiempo.

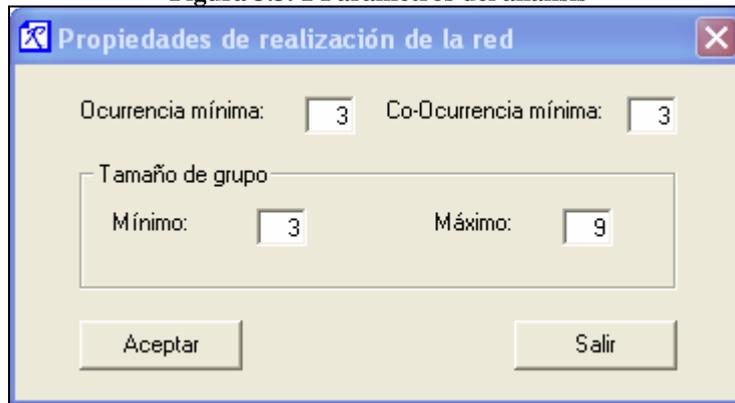
Un análisis por periodos cortos de tiempo y comparativo permitirá precisar la contribución al desarrollo en el área de las ciencias de la computación, pero antes de esto se revisó en detalle el algoritmo de palabras asociadas de esta herramienta, actividad que se describe en el siguiente apartado. El análisis por periodos cortos se describe en el apartado 3.4

### **3.3 Descripción del método de las palabras asociadas con algoritmo de agrupación sobre centros simples, utilizado por Redes 2005.**

La siguiente descripción del método se desarrollo con una muestra de las palabras de los títulos y los resúmenes de diez tesis del periodo 1996-2001.

#### **3.3.1 Selección de Parámetros para análisis en diez tesis.**

Con la experiencia de la literatura analizada (Ruiz, 2002), donde se elige el número de palabras con una ocurrencia y co-ocurrencia mínima de tres, y en el caso del tamaño mínimo y máximo de grupo se tomaron los valores que proporciona el software Redes 2005, que son 3 y 9 respectivamente, como los que se muestran en la Figura 3.3.1, los resultados para llegar a los agrupamientos de este periodo, se muestran a partir de la Figura 3.3.2 hasta la Figura 3.3.8.

**Figura 3.3. 1 Parámetros del análisis**

Fuente: Redes 2005 v.1.0.0.2008

En la Figura 3.3.2, se muestran los números de nodos o una estadística de palabras que resultaron de analizar las palabras que aparecen en los diez documentos de tesis del periodo 1996-2001 y que satisfacen los parámetros de la Figura 3.3.2.

**Figura 3.3. 2 Número de palabras que superan el umbral**

Fuente: Redes 2005 v.1.0.0.2008

En la Figura 3.3.3, se muestra un listado parcial de las 241 palabras que se identificaron, la lista se muestra en dos formatos, la primera ordenada por ocurrencia (frecuencia) y la segunda en orden alfabético.

**Figura 3.3. 3 Listado de palabras que superan el umbral**

Nodos. Ordenación por frecuencia			Nodos. Ordenación alfabética		
ID	Ocu.	Nombre	ID	Ocu.	Nombre
1	7	desarrollo	18	2	administracion
2	6	diseño	55	1	administracion-de-informacion
3	6	usuario	65	1	administracion-remota
4	5	herramienta	175	1	administrador-de-aplicaciones-distribuidas
5	5	informacion	27	2	algoritmo
6	5	sistema	137	1	algoritmos-geneticos
7	4	aplicacion	155	1	almacenar
8	4	comunicacion	178	1	alta-calidad
9	4	dato	182	1	ambiente-virtual
10	4	problema	221	1	ambientes-heterogeneos
11	3	java	39	1	analisis
12	3	lenguaje	53	1	analisis-espacial
13	3	modelo	226	1	analisis-y-diseño-objectory
14	3	programa	17	2	animacion
15	3	protocolo	49	1	api
16	3	tecnologia	7	4	aplicacion
17	2	animacion	85	1	aplicacion-en-internet
18	2	administracion	90	1	aplicaciones-distribuidas
19	2	backend	106	1	apuntador
20	2	bases-de-datos	115	1	arbol
21	2	cliente-servidor	131	1	arbol-de-busqueda
22	2	componente	138	1	archivo
23	2	estandar	149	1	arquitectura-abierta
24	2	extension	151	1	asignacion-de-fragmento
25	2	grafica	157	1	atributo
26	2	implementar	177	1	auditar
27	2	algoritmo	189	1	automatica
28	2	lenguaje-de-programacion	211	1	automatizacion
29	2	metodo	213	1	automatizacion-de-flujos
30	2	metodologia	218	1	automatizar
31	2	modulo	19	2	backend
32	2	proceso	38	1	base-de-datos-jerarquica

Fuente: Elaboración Propia con base de datos del CIC usando el software Redes 2005 v.1.0.0.2008

En la Figura 3.3.4, se muestra parcialmente la matriz de adyacencia resultante que incluye las 16 palabras que satisficieron el umbral, estas se encuentran en columnas y en renglones. Los valores que se localizan en la diagonal representan la ocurrencia o frecuencia de las correspondientes palabras y los valores que se localizan arriba de la diagonal representan la co-ocurrencia de las dos palabras que coinciden en esa columna-renglón.

**Figura 3.3. 4 Matriz de Adyacencia o co-ocurrencia con palabras.**

	1 desarrollo	2 diseño	3 usuario	4 herramien	5 informac	6 sistema	7 aplicacior	8 comunica	9 dato	10 problem	11 java	12 lenguaje	13 modelo	14
1 desarrollo	6	4	4	3	3	4	2	3	3	2	2	2	3	
2 diseño		5	4	2	2	3	2	3	2	2	2	3	2	
3 usuario			5	2	3	3	3	3	2	2	2	2	1	
4 herramien				4	3	4	2		1	2		2	2	
5 informac					4	3	2	1	1	2	1	1	2	
6 sistema						5	2	1	2	2	1	3	2	
7 aplicacior							3	1	1	2		1		
8 comunica								3	2	1	2	1	1	
9 dato									3	2	1	1	1	
10 problem										3		1	1	
11 java											2	1	1	
12 lenguaje												3	1	
13 modelo													3	
14 program.														
15 protocol														
16 tecnolog														

Fuente: Elaboración Propia con base de datos del CIC usando el software Redes 2005 v.1.0.0.2008

En la Figura 3.3.5, se muestra parcialmente la matriz normalizada resultante de las ya mostradas 16 palabras que satisficieron el umbral, estas se encuentran como el nombre tanto de las columnas como de los renglones. Los valores que se localizan arriba de la diagonal representan el índice de equivalencia, que se calcula mediante la siguiente formula:

$$e_{ij} = \frac{C_{ij}^2}{C_i C_j}$$

donde:

$e_{ij}$ : Índice de equivalencia o de asociación entre las palabras i y j.

$C_{ij}$ : Co-ocurrencia de las palabras i y j.

$C_i$ : Ocurrencia de la palabra i.

$C_j$ : Ocurrencia de la palabra j.

Ejemplo:

Sea la columna i = “herramienta”, sea el renglón j = “sistema”, los correspondientes valores de ocurrencia y co-ocurrencia de esta columna y renglón son  $C_i = 4$ ,  $C_j = 5$  y  $C_{ij} = 4$ , por lo tanto el valor del índice de equivalencia es:

$$e_{ij} = (C_{ij})^2 / C_i * C_j = 4^2 / 4 * 5 = 16 / 20 = 0.8$$

**Figura 3.3. 5 Matriz Normalizada con índice de equivalencia, periodo 1996-2001**

	1 desarrollo	2 diseño	3 usuario	4 herramien	5 informac	6 sistema	7 aplicacion	8 comunica	9 dato	10 problem	11 java	12 lenguaje	13 modelo	14
1 desarrollo	1	0,533	0,533	0,375	0,375	0,533	0	0,5	0,5	0	0	0	0	0,5
2 diseño		1	0,64	0	0	0,36	0	0,6	0	0	0	0	0,6	0
3 usuario			1	0	0,45	0,36	0,6	0,6	0	0	0	0	0	0
4 herramien				1	0,562	0,8	0	0	0	0	0	0	0	0
5 informac					1	0,45	0	0	0	0	0	0	0	0
6 sistema						1	0	0	0	0	0	0,6	0	0
7 aplicacion							1	0	0	0	0	0	0	0
8 comunica								1	0	0	0	0	0	0
9 dato									1	0	0	0	0	0
10 problem										1	0	0	0	0
11 java											1	0	0	0
12 lenguaje												1	0	0
13 modelo													1	0
14 program														1
15 protocol														
16 tecnolog														

Fuente: Elaboración Propia con base de datos del CIC usando el software Redes 2005 v.1.0.0.2008

En la Figura 3.3.6, se muestran los enlaces que se formaron por la co-ocurrencia de dos palabras y estos enlaces se encuentran ordenados por índice de equivalencia. Todos estos valores de índice de equivalencia se tomaron de la matriz normalizada.

**Figura 3.3. 6 Lista de enlaces ordenados por índice de equivalencia**

Rango	Selec.	ID1	Ocu	Nombre	ID2	Ocu	Nombre	Coocu	e
1	*	4	5	herramienta	6	5	sistema	4	0,800
2	*	2	6	diseño	3	6	usuario	4	0,640
3	*	2	6	diseño	8	4	comunicacion	3	0,600
4	*	2	6	diseño	12	3	lenguaje	3	0,600
5	*	3	6	usuario	7	4	aplicacion	3	0,600
6	*	3	6	usuario	8	4	comunicacion	3	0,600
7	*	6	5	sistema	12	3	lenguaje	3	0,600
8	*	4	5	herramienta	5	5	informacion	3	0,562
9	*	1	7	desarrollo	2	6	diseño	4	0,533
10	*	1	7	desarrollo	3	6	usuario	4	0,533
11	*	1	7	desarrollo	6	5	sistema	4	0,533
12	*	1	7	desarrollo	8	4	comunicacion	3	0,500
13	*	1	7	desarrollo	9	4	dato	3	0,500
14	*	1	7	desarrollo	13	3	modelo	3	0,500
15	*	1	7	desarrollo	16	3	tecnologia	3	0,500
16	*	3	6	usuario	5	5	informacion	3	0,450
17	*	5	5	informacion	6	5	sistema	3	0,450
18	*	1	7	desarrollo	4	5	herramienta	3	0,375
19	*	1	7	desarrollo	5	5	informacion	3	0,375
20	*	2	6	diseño	6	5	sistema	3	0,360
21	*	3	6	usuario	6	5	sistema	3	0,360

Fuente: Elaboración Propia con base de datos del CIC usando el software Redes 2005 v.1.0.0.2008

Para formar los agrupamientos de los temas o palabras centrales, a la lista de enlaces de la Figura 3.3.6 se le aplico el *algoritmo de agrupación por centros simples* (<http://www.ugr.es/~rruizb/cognosfera/index.htm>), el cual se auxilio de la siguiente matriz en Excel. Figura 3.3.7

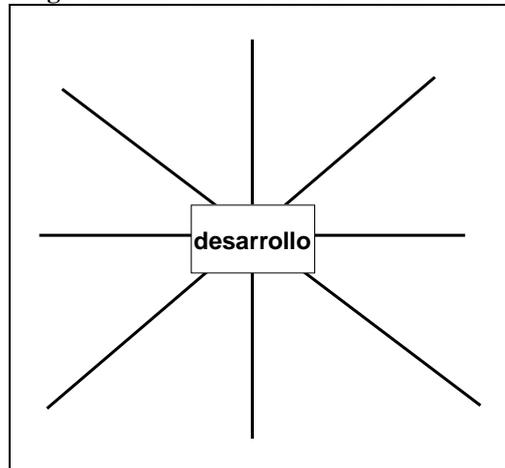
**Figura 3.3. 7 Matriz con resultado parcial del algoritmo de agrupación por centros simples**

Rango	id1	id2	desarrollo	diseño	usuario	herramienta	informacion	sistema	aplicacion	comunicacion	dato	problema	java	lenlenguaje	modelo	programa	protocolo	tecnologia
1	herramienta	sistema			1	1												
2	diseño	usuario		1	1													
3	diseño	comunicación		2					1									
4	diseño	lenguaje		3									1					
5	usuario	aplicación			2			1										
6	usuario	comunicación			3				2									
7	sistema	lenguaje					2						2					
8	herramienta	informacion				2	1											
9	desarrollo	diseño	1	4														
10	desarrollo	usuario	2		4													
11	desarrollo	sistema	3				3											
12	desarrollo	comunicación	4						3									
13	desarrollo	dato	5								1							
14	desarrollo	modelo	6											1				
15	desarrollo	tecnologia	7															
16	usuario	informacion			5	2												
17	informacion	sistema				3	4											
18	desarrollo	herramienta	8			3												
19	desarrollo	informacion	9			4												
20	diseño	sistema		5				5										
21	usuario	sistema			6			6										

Fuente: Elaboración Propia con base de datos del CIC.2008

En la Figura 3.3.8 muestra como el algoritmo de agrupación de centros simples va generando el agrupamiento, en este caso alrededor de la palabra central “desarrollo” que se encuentra en el renglón 18 de la Figura 3.3.7, quien alcanzo el tamaño máximo menos uno del grupo a formar, parámetro especificado en la Figura 3.3.1

Figura 3.3. 8 Palabra central “desarrollo”



Fuente: Elaboración Propia a partir de datos del CIC.2008

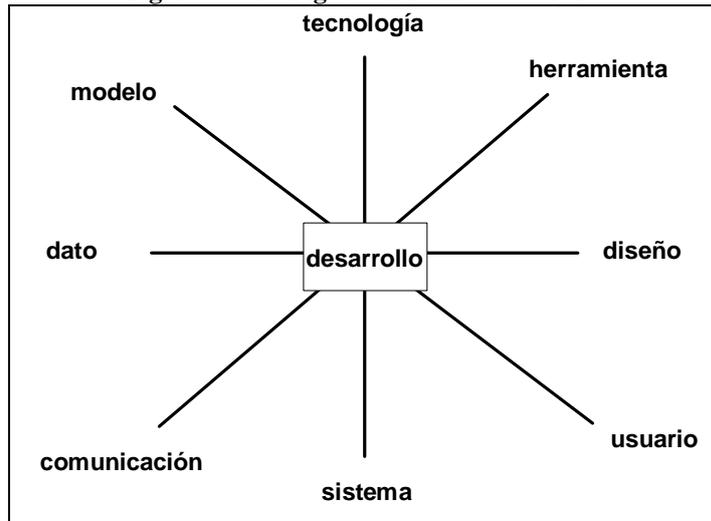
Para etiquetar los enlaces de la Figura 3.3.8 con sus correspondientes descriptores, se hace uso de la matriz de la Figura 3.3.9 donde se toman las palabras de los primeros ocho enlaces de arriba hacia abajo de la palabra “desarrollo”, en este caso, diseño, usuario, sistema, comunicación, dato, modelo, tecnología y herramienta, con lo que se obtiene la Figura 3.3.10

Figura 3.3. 9 Matriz con resultado parcial del algoritmo de agrupación por centros simples.

Rango	id1	id2	desarrollo	diseño	usuario	herramienta	información	sistema	aplicación	comunicación	dato	problema	java	lenguaje	modelo	programa	protocolo	tecnología
1	herramienta	sistema			1		1											
2	diseño	usuario		1	1													
3	diseño	comunicación		2					1									
4	diseño	lenguaje		3										1				
5	usuario	aplicación			2			1										
6	usuario	comunicación			3				2									
7	sistema	lenguaje						2						2				
8	herramienta	informacion				2	1											
9	desarrollo	diseño	1	4														
10	desarrollo	usuario	2		4													
11	desarrollo	sistema	3				3											
12	desarrollo	comunicación	4						3									
13	desarrollo	dato	5							1								
14	desarrollo	modelo	6											1				
15	desarrollo	tecnología	7														1	
16	usuario	informacion			5		2											
17	informacion	sistema					3	4										
18	desarrollo	herramienta	8			3												
19	desarrollo	informacion	9				4											
20	diseño	sistema		5				5										
21	usuario	sistema			6			6										

Fuente: Elaboración Propia con base de datos del CIC.2008

Figura 3.3. 10 Figura en forma de estrella.



Fuente: Elaboración Propia con base de datos del CIC.2008

El siguiente paso es adicionar los enlaces entre los descriptores del tema “desarrollo”. Por cada uno de los descriptores se adicionan los enlaces con los otros descriptores del grupo “desarrollo”.

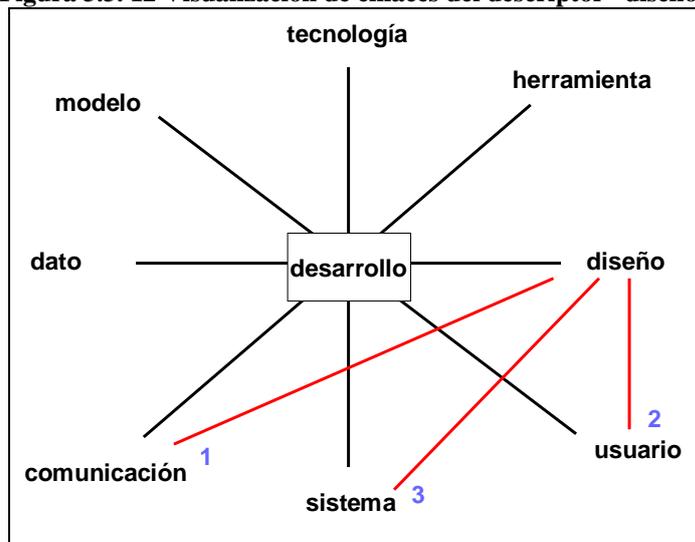
Por ejemplo. En la Figura 3.3.11 se muestra los enlaces del primer descriptor tomado en cuenta, en este caso “diseño”. Los correspondientes enlaces se encuentran en los renglones 2, 3, 4, 9 y 20, que corresponden a las palabras “usuario”, “comunicación”, “lenguaje”, “desarrollo” y “sistema”. Los enlaces que se dibujaran en el grafo de la Figura 3.3.9 serán con las palabras que pertenecen al grupo de desarrollo, es decir con “usuario”, “comunicación”, “desarrollo” y “sistema”, esta adición de enlaces se muestra en la Figura 3.3.12.

Figura 3.3. 11 Señalamiento de las palabras de los enlaces del descriptor “diseño”

Rango	id1	id2	desarrollo	diseño	usuario	herramienta	información	sistema	aplicación	comunicación	dato	problema	java	lenguaje	modelo	programa	protocolo	tecnología	
1	herramienta	sistema				1		1											
2	diseño	usuario	1	1															
3	diseño	comunicación	2						1										
4	diseño	lenguaje	3											1					
5	usuario	aplicación			2			1											
6	usuario	comunicación			3				2										
7	sistema	lenguaje					2							2					
8	herramienta	información				2	1												
9	desarrollo	diseño	1	4															
10	desarrollo	usuario	2		4														
11	desarrollo	sistema	3					3											
12	desarrollo	comunicación	4						3										
13	desarrollo	dato	5							1									
14	desarrollo	modelo	6											1					
15	desarrollo	tecnología	7															1	
16	usuario	información			5		2												
17	información	sistema					3	4											
18	desarrollo	herramienta	8			3													
19	desarrollo	información	9				4												
20	diseño	sistema		5				5											
21	usuario	sistema			6			6											

Fuente: Elaboración Propia con base de datos del CIC.2008

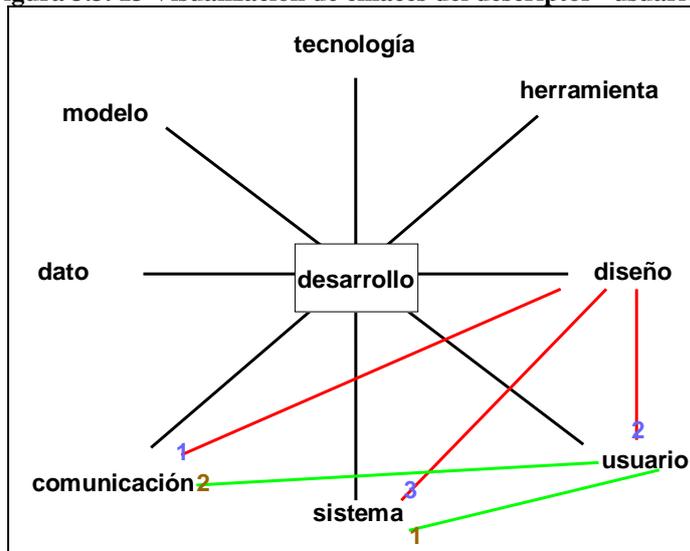
Figura 3.3. 12 Visualización de enlaces del descriptor “diseño”.



Fuente: Elaboración Propia con base de datos del CIC.2008

Repetiendo el proceso anterior con el descriptor “usuario”, se obtiene el grafo de la Figura 3.3.13

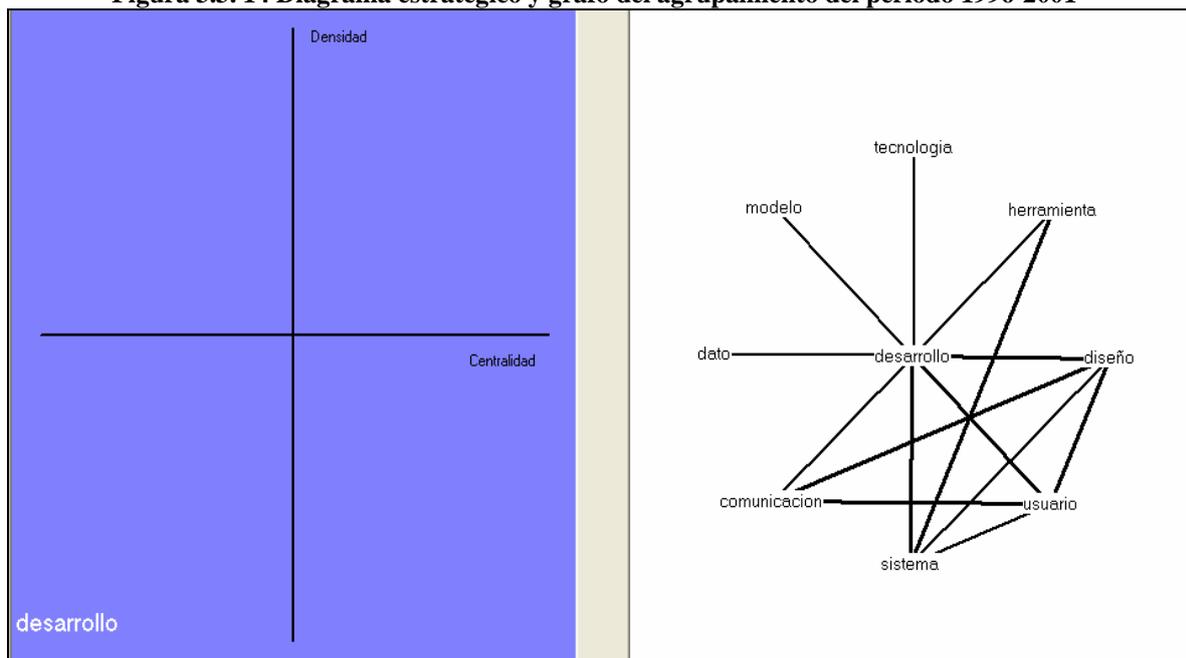
Figura 3.3. 13 Visualización de enlaces del descriptor “usuario”.



Fuente: Elaboración Propia con base de datos del CIC.2008

Al repetir este proceso con los restantes seis descriptores del grupo se obtiene la Figura 3.3.14

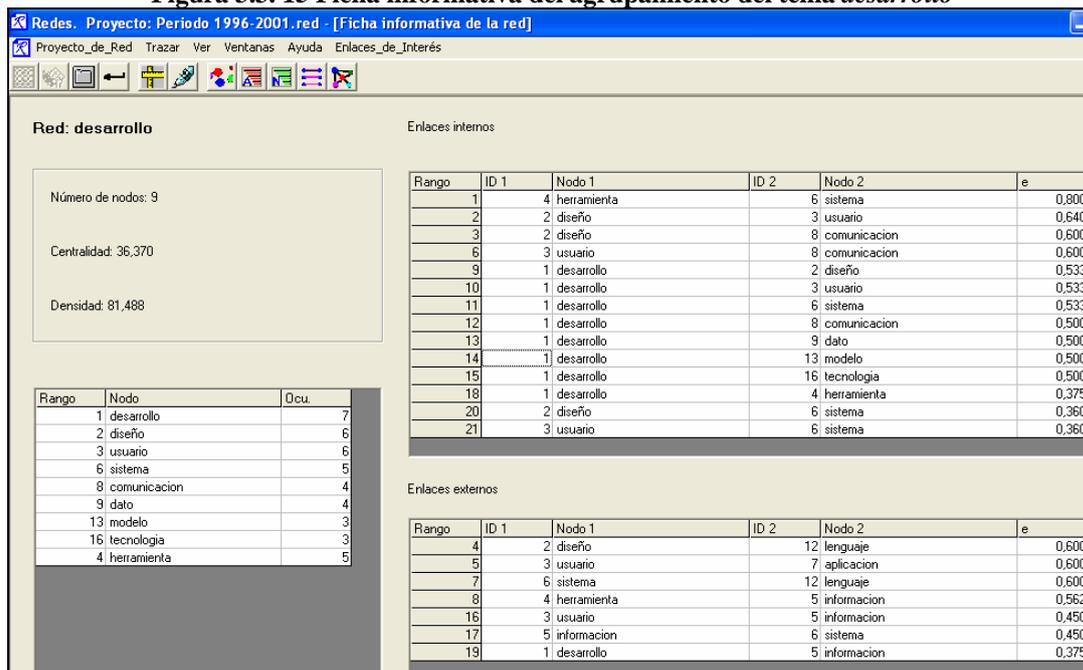
Figura 3.3. 14 Diagrama estratégico y grafo del agrupamiento del periodo 1996-2001



Fuente: Elaboración Propia con base de datos del CIC usando el software Redes 2005 v.1.0.0.2008

Los enlaces que se muestran en la Figura 3.3.14 y que se tomaron de la Figura 3.3.6 se conocen como *enlaces internos*. Cualquier otro enlace de un descriptor o palabra que pertenezca a este grupo y que se una a otro descriptor o palabra de otro grupo se le llama *enlace externo*. Figura 3.3.15

Figura 3.3. 15 Ficha informativa del agrupamiento del tema *desarrollo*



Fuente: Elaboración Propia con base de datos del CIC usando el software Redes 2005 v.1.0.0.2008

Una vez que se identificaron los temas o palabras centrales, los descriptores que definen a este tema o palabras centrales y los enlaces que los unen. Para poder ubicar en el diagrama estratégico, es necesario calcular el *índice de densidad*, el *índice de centralidad*, el *rango de densidad* y el *rango de centralidad*.

- El cálculo del índice de densidad es de la siguiente forma:

$$d = 100 \frac{\sum_{i=1}^L e_i}{P}$$

donde:

ei: Índice de equivalencia del enlace interno i

L: Número de enlaces internos del tema.

P: Número de palabras del tema.

Ejemplo:

Vamos a calcular el índice de densidad del tema *“desarrollo”*. Las variables que intervienen en este cálculo con su valor son:

$$L = 14$$

$$P = 9$$

$e_i$  = valores de la columna etiquetada con la letra “e” de la Figura 4.4.15, substituyendo en la siguiente formula

$$d = 100 \frac{\sum_{i=1}^L e_i}{P}$$

obtenemos el siguiente calculo

$$d = (0.80 + 0.64 + 0.60 + 0.60 + 0.533 + 0.533 + 0.533 + 0.50 + 0.50 + 0.50 + 0.50 + 0.375 + 0.36 + 0.36/9) \\ * 100 = 81.488$$

Para obtener el rango de densidad:

$$r_d = \frac{\text{RANGO}}{N}$$

donde:

rango: Rango del tema según su densidad

N: Número de temas de la red.

Ejemplo:

$$rd = 1/1 = 1$$

- El cálculo del índice de centralidad es de la siguiente forma:

$$c = 10 \sum_{j=1}^T e_j$$

donde:

$e_j$ : Índice de equivalencia del enlace externo j

T: Número total de enlaces externos.

Ejemplo:

Vamos a calcular el índice de centralidad del tema “desarrollo”. Las variables que intervienen en este cálculo con su valor son:

$T = 7$

$e_i$  = valores de la columna etiquetada con la letra “e” de la Figura 4.4.15, substituyendo en la siguiente formula

$$c = 10 \sum_{j=1}^T e_j$$

obtenemos el siguiente calculo

$$c = (0.60+0.60+0.60+0.562+0.450+0.450+0.375)*10= 36.370$$

- Para obtener el rango de centralidad:

$$r_c = \frac{\text{rango}^r}{N}$$

donde:

rango: Rango del tema según su centralidad

N: Número de temas de la red

$$r_c = 1/1=1$$

### 3.3.1.1 Interpretación del tema *desarrollo* en el diagrama estratégico.

En la Figura 3.3.14 se muestra en la izquierda el diagrama estratégico que corresponde a las diez tesis del periodo 1996-2001, en él se muestra un solo tema o palabra central “*desarrollo*” con una centralidad y densidad baja , lo que explica que es un tema que se encuentra en decadencia o un tema naciente, en la derecha se muestran los descriptores que se encuentran relacionados con el tema o palabra central siendo este “*desarrollo*” y que son *diseño, usuario, sistema, comunicación, modelo, tecnología, herramienta e información*. Con esto se ejemplifico a detalle como se construyen los agrupamientos.

### 3.4 Análisis de tópicos de las ciencias de la computación en documentos de tesis.

Para identificar y describir las relaciones o los agrupamientos en los tópicos o conceptos de las ciencias de la computación se llevo a cabo el análisis en las palabras tanto del resumen como del título de 90 tesis, divididas en tres periodos de tiempo, siendo estos 1996-2001; 2002-2004; 2005-2007; como ya se mencionó, los tópicos que se estudian en

los documentos de los temas de tesis se agrupan y describen los temas relevantes en cada uno de los periodos.

### 3.4.1 Análisis del Período 1996-2001

Para describir en forma resumida el proceso de análisis del periodo 1996-2001, con la herramienta Redes 2005, en la Tabla 3.4.1.1 se muestran los parámetros y los resultados de este proceso.

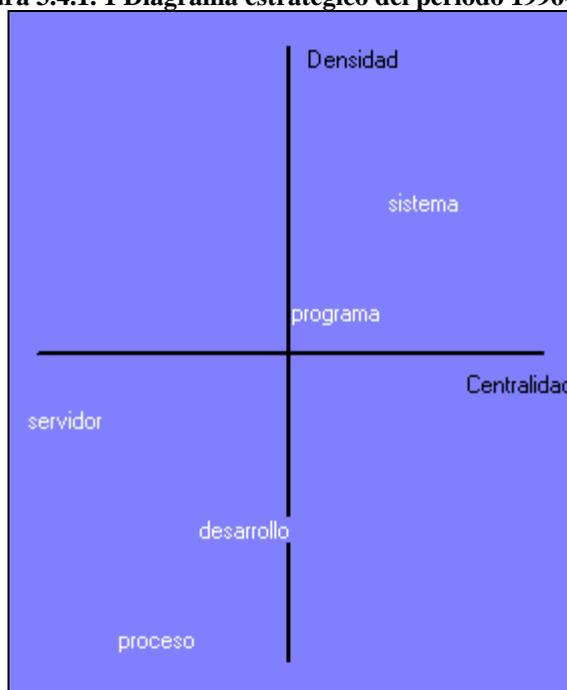
**Tabla 3.4.1. 1 Parámetros y Resultados de la ejecución con Redes 2005 en el periodo 1996-2001**

Parámetros	Análisis de las palabras	Tema	Nodos	Enlaces Internos	Densidad	Rd	Enlaces Externos	Centralidad	Rc
Ocurrencia mínima =3	nodos identificados= 662	sistema	9	24	66.022	5/5	50	179.390	5/5
Co-ocurrencia mínima=3	Nodos que superan el umbral= 68	programa	9	18	61.922	4/5	50	162.100	4/5
Tamaño del grupo mínimo=3	Nodos descartados= 594	servidor	2	20	43.733	3/5	0	0.000	3/5
Tamaño del grupo máximo=9		desarrollo	9	8	22.044	2/5	14	29.940	1/5
		proceso	5	4	20.340	1/5	0	0.000	2/5

**Fuente: Elaboración Propia con base de datos del CIC.2008**

La Figura 3.4.1.1 muestra el diagrama estratégico con los cinco temas que resultaron del análisis en el periodo 1996-2001 y que se indican en la Tabla 3.4.1.1.

**Figura 3.4.1. 1 Diagrama estratégico del periodo 1996-2001**



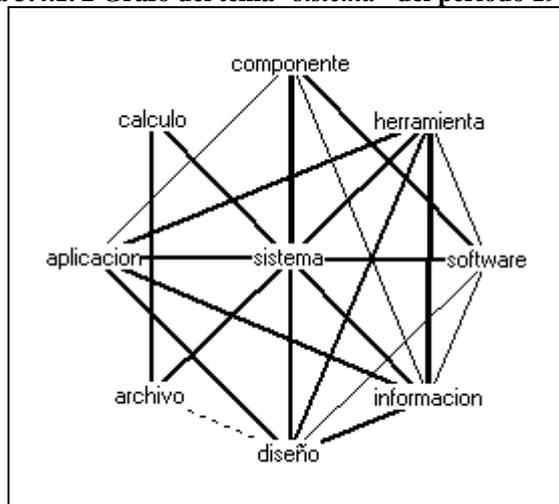
**Fuente: Elaboración Propia con base de datos del CIC usando el software Redes 2005 v.1.0.0.2008**

De acuerdo a la forma en que se construyen los diagramas estratégicos, una primera interpretación de los temas en este diagrama es la siguiente:

- a) En el primer cuadrante se sitúan los temas que generalmente se encuentran muy desarrollados y juegan un papel muy importante en el campo de estudio, es este caso los temas “*sistema*” y “*programa*” en el campo de las ciencias de la computación en el CIC.
- b) En el cuarto cuadrante se encuentran los temas que aparecen por primera vez y que en muchos casos terminan aquí por desaparecer o son susceptibles de convertirse en emergentes, en este caso los temas “*servidor*”, “*desarrollo*” y “*proceso*”.
- c) El diagrama estratégico de la Figura 3.4.1.1 muestra una situación que de acuerdo a Callon, Courtail y Penan (1995) se ubican en la categoría 1 y que significa que son temas que se encuentran bien relacionados y desarrollados.

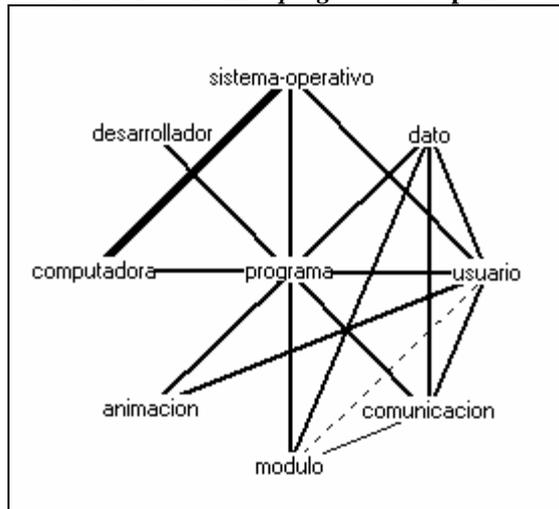
De la Figura 3.4.1.2 a la Figura 3.4.1.6 se muestra el detalle de cada uno de los temas y sus descriptores que se localizaron en el análisis de este periodo.

**Figura 3.4.1. 2 Grafo del tema “*sistema*” del periodo 1996-2001**



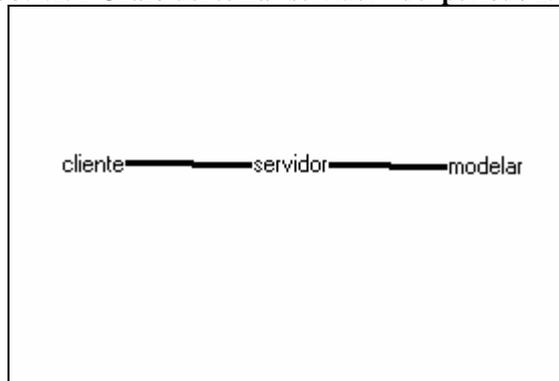
**Fuente: Elaboración Propia con base de datos del CIC usando el software Redes 2005 v.1.0.0.2008**

**Figura 3.4.1. 1 Grafo del tema “programa” del periodo 1996-2001**



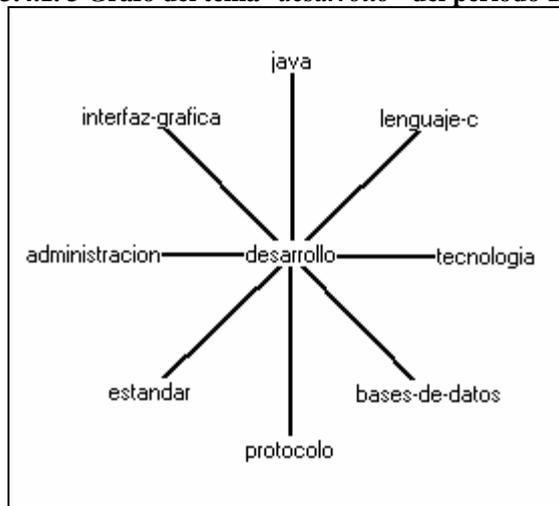
Fuente: Elaboración Propia con base de datos del CIC usando el software Redes 2005 v.1.0.0.2008

**Figura 3.4.1. 2 Grafo del tema “servidor” del periodo 1996-2001**



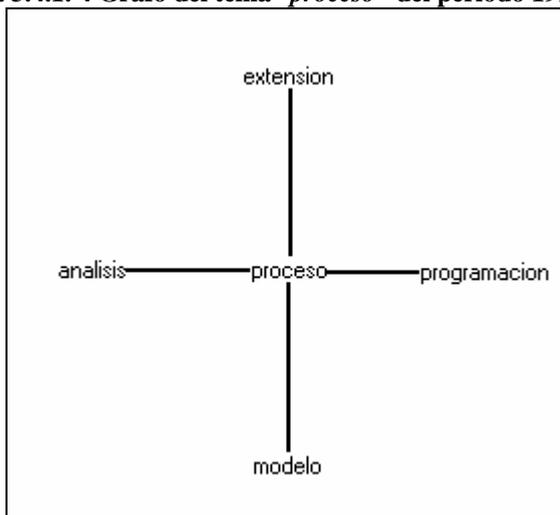
Fuente: Elaboración Propia con base de datos del CIC usando el software Redes 2005 v.1.0.0.2008

**Figura 3.4.1. 3 Grafo del tema “desarrollo” del periodo 1996-2001**



Fuente: Elaboración Propia con base de datos del CIC usando el software Redes 2005 v.1.0.0.2008

**Figura 3.4.1. 4 Grafo del tema “proceso” del periodo 1996-2001**



**Fuente: Elaboración Propia con base de datos del CIC usando el software Redes 2005 v.1.0.0.2008**

### 3.4.2 Análisis del Período 2002-2004

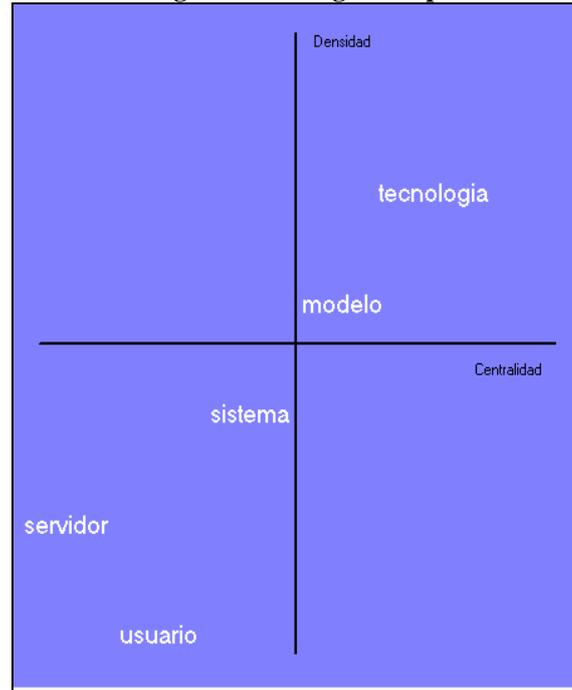
Para describir en forma resumida el proceso de análisis del periodo 2002-2004, con la herramienta Redes 2005, en la Tabla 3.4.2.1 se muestran los parámetros y los resultados de este proceso.

**Tabla 3.4.2. 2 Parámetros y Resultados de la ejecución con Redes 2005 en el periodo 2002-2004**

Parámetros	Construir la red	Temas	Nodos	Enlaces Internos	Densidad	Rd	Enlaces Externos	Centralidad	Rc
Ocurrencia mínima =3	Identificación de nodos= 729	tecnología	9	24	87.256	5/5	50	197.760	5/5
Co-ocurrencia mínima=3	Nodos que superan el umbral= 63	modelo	9	30	82.300	4/5	50	126.200	4/5
Tamaño del grupo mínimo=3	Nodos descartados= 666	sistema	9	20	69.011	3/5	34	90.320	3/5
Tamaño del grupo máximo=9		servidor	4	4	47.800	2/5	1	3.210	1/5
		usuario	5	4	19.500	1/5	4	22.120	2/5

**Fuente: Elaboración Propia con base de datos del CIC.2008**

La Figura 3.4.2.1 muestra el diagrama estratégico con los cinco temas que resultaron del análisis en el periodo 2002-2004 y que se indican en la Tabla 3.4.2.1.

**Figura 3.4.2. 5 Diagrama estratégico del periodo 2002-2004**

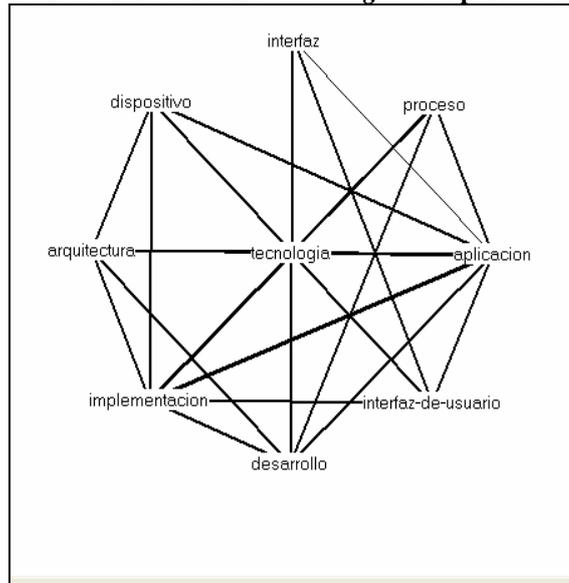
**Fuente: Elaboración Propia con base de datos del CIC usando el software Redes 2005 v.1.0.0.2008**

De acuerdo a la forma en que se construyen los diagramas estratégicos, una primera interpretación de los temas en este diagrama es la siguiente:

- d) En el primer cuadrante se sitúan los temas que generalmente se encuentran muy desarrollados y juegan un papel muy importante en el campo de estudio, es este caso los temas “*tecnología*” y “*modelo*” en el campo de las ciencias de la computación en el CIC.
- e) En el cuarto cuadrante se encuentran los temas que aparecen por primera vez y que en muchos casos terminan aquí por desaparecer o son susceptibles de convertirse en emergentes, en este caso los temas “*sistema*”, “*servidor*” y “*usuario*”.
- f) El diagrama estratégico de la Figura 3.4.2.1 muestra una situación que de acuerdo a Callon, Courtail y Penan (1995) se ubican en la categoría 1 y que significa que son temas que se encuentran bien relacionados y desarrollados.

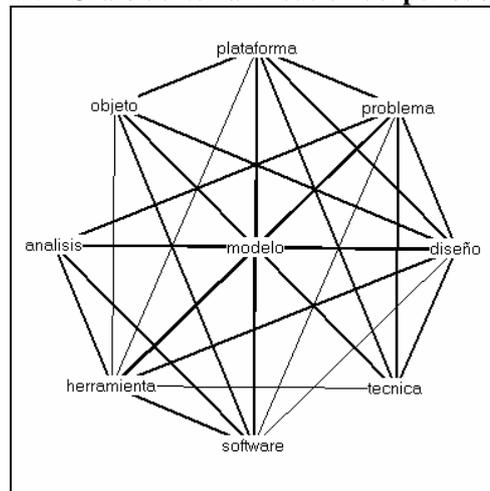
De la Figura 3.4.2.2 a la Figura 3.4.2.6 se muestra el detalle de cada uno de los temas que se localizaron en el análisis de este periodo.

**Figura 3.4.2. 6 Grafo del tema “tecnología” del periodo 2002-2004**



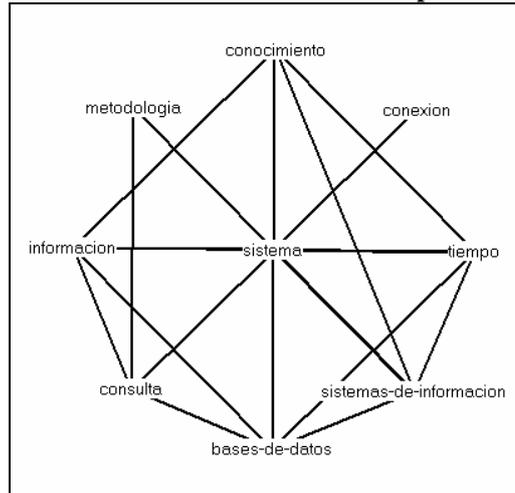
**Fuente: Elaboración Propia con base de datos del CIC usando el software Redes 2005 v.1.0.0.2008**

**Figura 3.4.2. 7 Grafo del tema “modelo” del periodo 2002-2004**



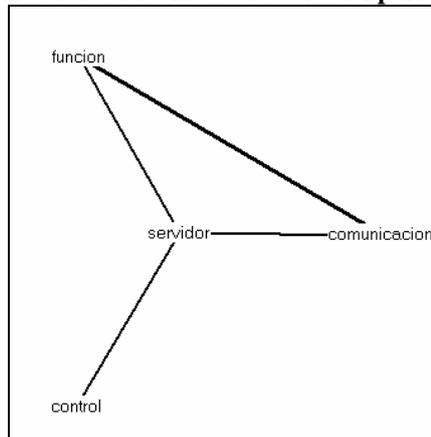
**Fuente: Elaboración Propia con base de datos del CIC usando el software Redes 2005 v.1.0.0.2008**

**Figura 3.4.2. 8 Grafo del tema “sistema” del periodo 2002-2004**



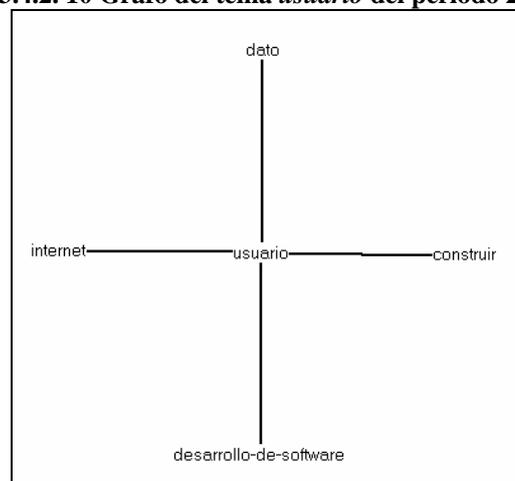
**Fuente: Elaboración Propia con base de datos del CIC usando el software Redes 2005 v.1.0.0.2008**

**Figura 3.4.2. 9 Grafo del tema “servidor” del periodo 2002-2004**



**Fuente: Elaboración Propia con base de datos del CIC usando el software Redes 2005 v.1.0.0.2008**

**Figura 3.4.2. 10 Grafo del tema *usuario* del periodo 2002-2004**



**Fuente: Elaboración Propia con base de datos del CIC usando el software Redes 2005 v.1.0.0.2008**

### 3.4.3 Análisis del Período 2005-2007

Para describir la forma resumida del proceso de análisis del periodo 2005-2007, con la herramienta Redes 2005, en la Tabla 3.4.3.1 se muestran los parámetros y los resultados de este proceso.

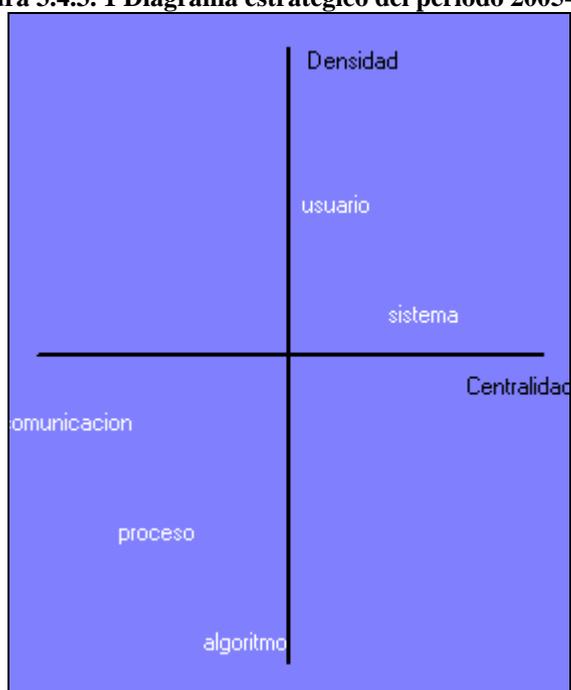
**Tabla 3.4.3. 1 Parámetros y Resultados de la ejecución con Redes 2005 en el periodo 2005-2007.**

Parámetros	Construir la red	Temas	Nodos	Enlaces Internos	Densidad	Rd	Enlaces Externos	Centralidad	Rc
Ocurrencia mínima =3	Identificación de nodos= 728	Usuario	9	18	48.478	5/5	50	171.820	5/5
Co-ocurrencia mínima=3	Nodos que superan el umbral=74	Sistema	9	18	48.478	4/5	50	171.820	4/5
Tamaño del grupo mínimo=3	Nodos descartados= 654	Comunicación	3	2	23.200	3/5	1	2.500	1/5
Tamaño del grupo máximo=9		Proceso	8	9	22.175	2/5	6	18.380	3/5
		Algoritmo	9	10	21.067	1/5	11	21.067	2/5

**Fuente: Elaboración Propia con base de datos del CIC.2008**

La Figura 3.4.3.1 muestra el diagrama estratégico con los seis temas que resultaron del análisis en el periodo 2005-2007 y que se indican en la Tabla 3.4.3.1.

**Figura 3.4.3. 1 Diagrama estratégico del periodo 2005-2007**



**Fuente: Elaboración Propia con base de datos del CIC usando el software Redes 2005 v.1.0.0.2008**

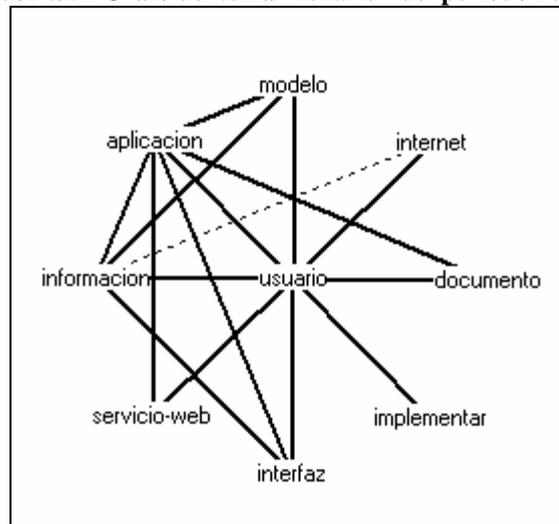
En forma similar de cómo se llevo a cabo la interpretación de los resultados del análisis del periodo 2002-2004 y como ya se mencionó por la forma en que se construyen los

diagramas estratégicos, una primera interpretación de los temas en este diagrama es la siguiente:

- a) Como ya se menciona en el primer cuadrante se sitúan los temas que generalmente se encuentran muy desarrollados y juegan un papel muy importante en el campo de estudio, es este caso son los temas “*usuario*” y “*sistema*”.
- b) En el cuarto cuadrante se encuentran los temas que aparecen por primera vez y que en muchos casos terminan aquí por desaparecer, en este caso los temas “*comunicación*”, “*proceso*” y “*algoritmo*”.
- c) En el diagrama estratégico de la Figura 3.4.3.1 muestra una situación similar al diagrama estratégico de la Figura 3.4.2.1, es decir se tiene nuevamente la categoría 1, que significa que son temas que se encuentran bien relacionados y desarrollados y que tienen contacto con un conjunto de temas poco desarrollados.

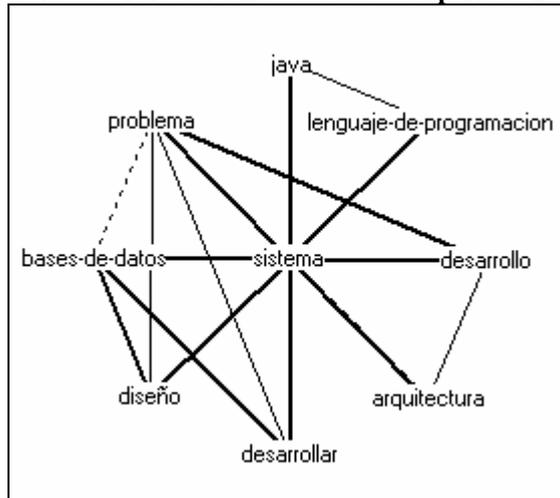
De la Figura 3.4.3.2 a la Figura 3.4.3.6 se muestra el detalle de cada uno de los temas que se localizaron en el análisis de este periodo.

**Figura 3.4.3. 2 Grafo del tema “*usuario*” del periodo 2005-2007**



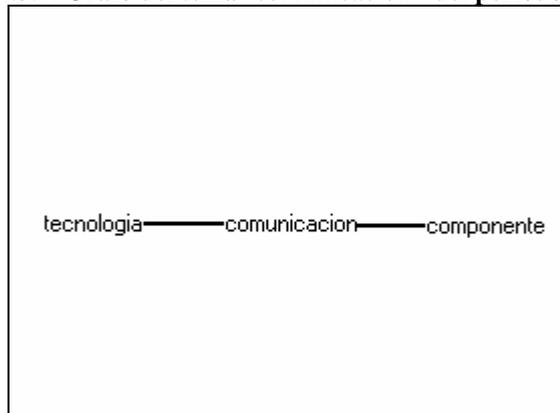
**Fuente: Elaboración Propia con base de datos del CIC usando el software Redes 2005 v.1.0.0.2008**

**Figura 3.4.3. 3 Grafo del tema “sistema” del periodo 2005-2007**



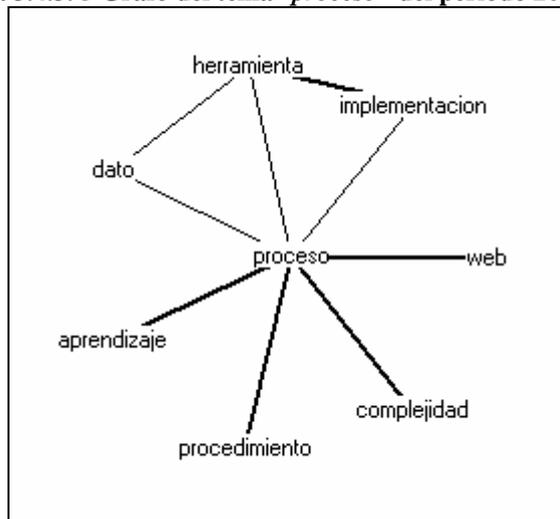
**Fuente: Elaboración Propia con base de datos del CIC usando el software Redes 2005 v.1.0.0.2008**

**Figura 3.4.3. 4 Grafo del tema “comunicación” del periodo 2005-2007**

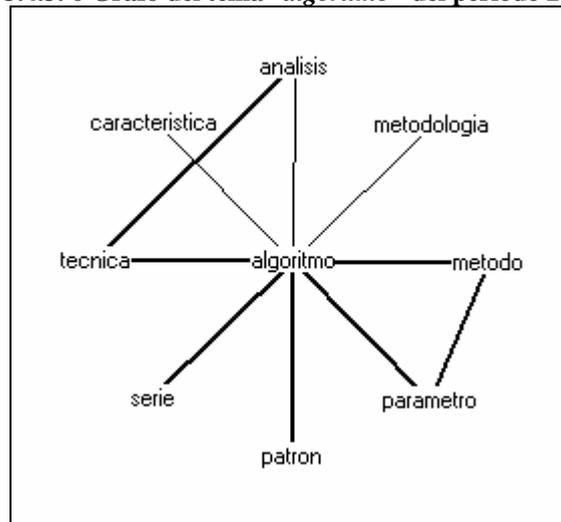


**Fuente: Elaboración Propia con base de datos del CIC usando el software Redes 2005 v.1.0.0.2008**

**Figura 3.4.3. 5 Grafo del tema “proceso” del periodo 2005-2007**



**Fuente: Elaboración Propia con base de datos del CIC usando el software Redes 2005 v.1.0.0.2008**

Figura 3.4.3. 6 Grafo del tema “*algoritmo*” del periodo 2005-2007

Fuente: Elaboración Propia con base de datos del CIC usando el software Redes 2005 v.1.0.0.2008

# Capítulo 4

## Búsqueda de patrones y tendencias en los grafos y diagramas estratégicos

En cualquier campo de conocimiento es importante conocer el ciclo de vida de sus áreas que lo conforman, el saber el ciclo de vida de un área, ciclo que generalmente se relaciona con la producción científica de esta, podría ayudar a la toma de decisiones como las siguientes:

- a) Aplicar las políticas de apoyo a un área que va emergiendo como futura área en gran desarrollo y por consiguiente con gran productividad.
- b) Que los alumnos del campo de conocimiento, seleccionen esta área como uno de sus temas de tesis.
- c) Monitorear el progreso del área y cuando se encuentre en una situación de declive, enfatizar su atención en ella.

### 4.1 Metodología de búsqueda de patrones y tendencias

Los pasos para localizar los patrones y tendencias en un área de conocimiento es bajo la siguientes premisas, un tópico de computación se muestra en el diagrama estratégico como un tema.

- a) El tema a buscar debe tener un comportamiento a través del tiempo en las siguientes secuencias de cuadrantes.
  - a1) 4, 3 y 1
  - a2) 4, 2 y 1
  - a3) 4 y 1

- b) El tema que tenga alguno de los comportamientos anteriores deberá de haber sido un descriptor de otro tema, esta búsqueda se realiza en los grafos, de igual forma como se hizo mención en el párrafo anterior, este análisis es a través del tiempo.
- c) Además el descriptor deberá tener la característica de que incremente sus enlaces internos o enlaces externos, para que después, bajo los valores de parámetros este descriptor se forme como el representante de un tema con un comportamiento del inciso a.

#### 4.2 Análisis de los diagramas estratégicos y grafos en los periodos

Para poder aplicar la metodología, primero se hace un resumen del comportamiento de los tópicos a través de los cuadrantes en los diagramas estratégicos de los diferentes periodos.

Tabla 4.2.1

Tabla 4.2. 1 Ubicación de un tema en diagramas estratégicos.

<b>Tema o descriptor</b>	<b>Periodo 1 1996-2001</b>	<b>Periodo 2 2002-2004</b>	<b>Periodo 3 2005-2007</b>
	<b>Cuadrantes</b>		
<b>Sistema</b>	<b>1</b>	<b>4</b>	<b>1</b>
<b>Programa</b>	<b>1</b>	-	-
<b>Servidor</b>	<b>4</b>	<b>4</b>	-
<b>Desarrollo</b>	<b>4</b>	-	-
<b>Proceso</b>	<b>4</b>	-	<b>4</b>
<b>Tecnología</b>	-	<b>1</b>	-
<b>Modelo</b>	-	<b>1</b>	-
<b>Usuario</b>	-	<b>4</b>	<b>1</b>
<b>Comunicación</b>	-	-	<b>4</b>
<b>Algoritmo</b>	-	-	<b>4</b>

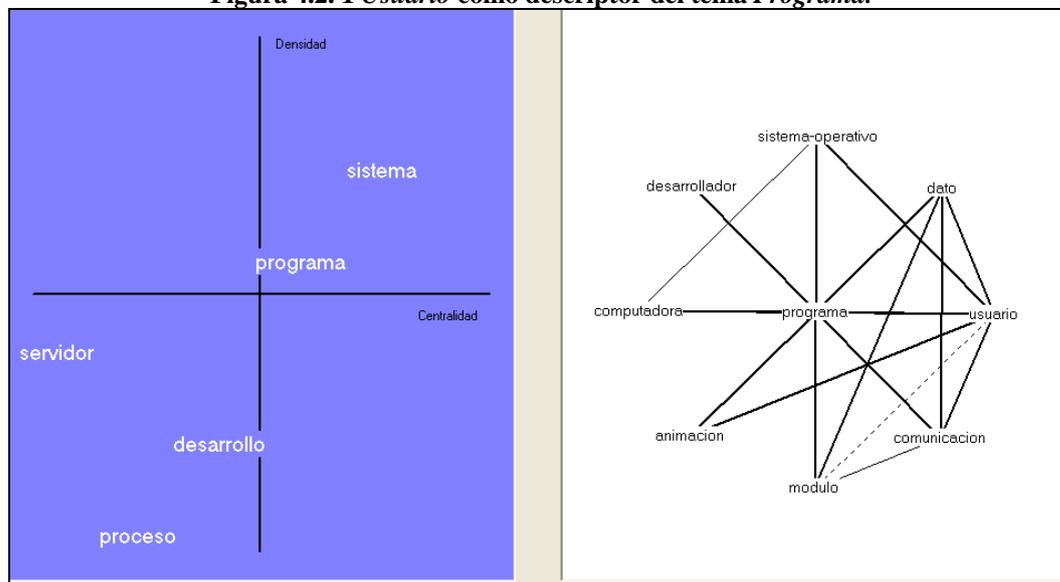
Fuente: Elaboración Propia con base de datos del CIC.2008

Al revisar la Tabla 4.2.1, el tema que puede tener un comportamiento de los descritos en el inciso a) del punto 4.1, es el tema *usuario*, el cual tiene la secuencia a3)

Como lo marca la metodología en el inciso b), se buscó que este tema sea un descriptor de otro tema en un periodo anterior, para lo cual se reviso los grafos del periodo 1996-2001.

La Figura 4.2.1 muestra el grafo del tema *programa* que corresponde al periodo 1996-2001, donde aparece como descriptor el tópico *usuario*.

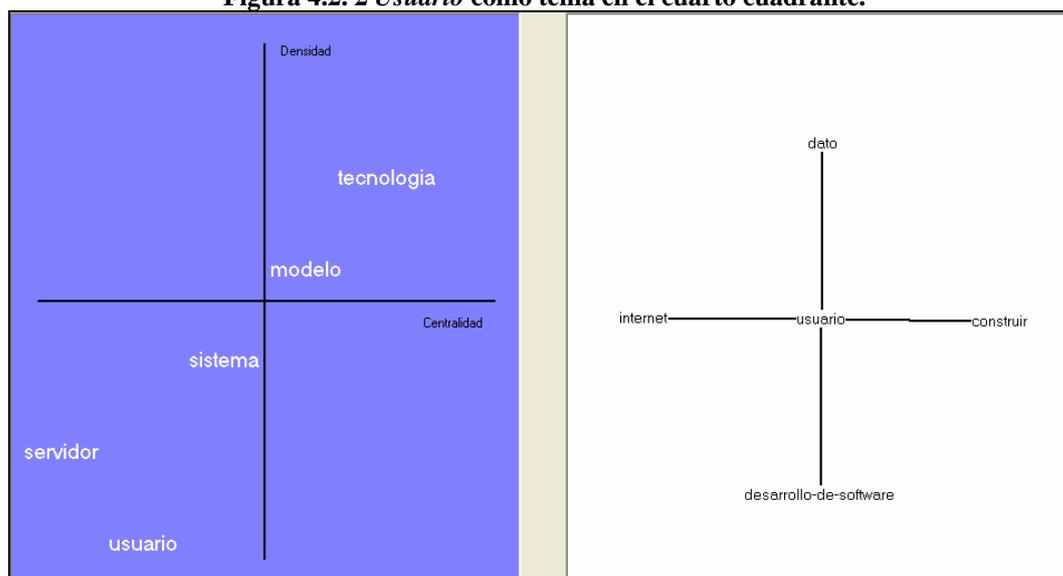
**Figura 4.2.1 Usuario como descriptor del tema Programa.**



Fuente: Elaboración Propia con base de datos del CIC usando el software Redes 2005 v.1.0.0.2008

Como se observa en la Figura 4.2.1.1 el descriptor usuario cumple con el inciso c) del punto 3.2, su número de enlaces internos se incrementó y es casi igual al número de enlaces del tema central, es decir próximo a convertirse en un tema en otro periodo, como se muestra en la Figura 4.2.2 y que corresponde al periodo 2002-2004, situándose en el cuarto cuadrante.

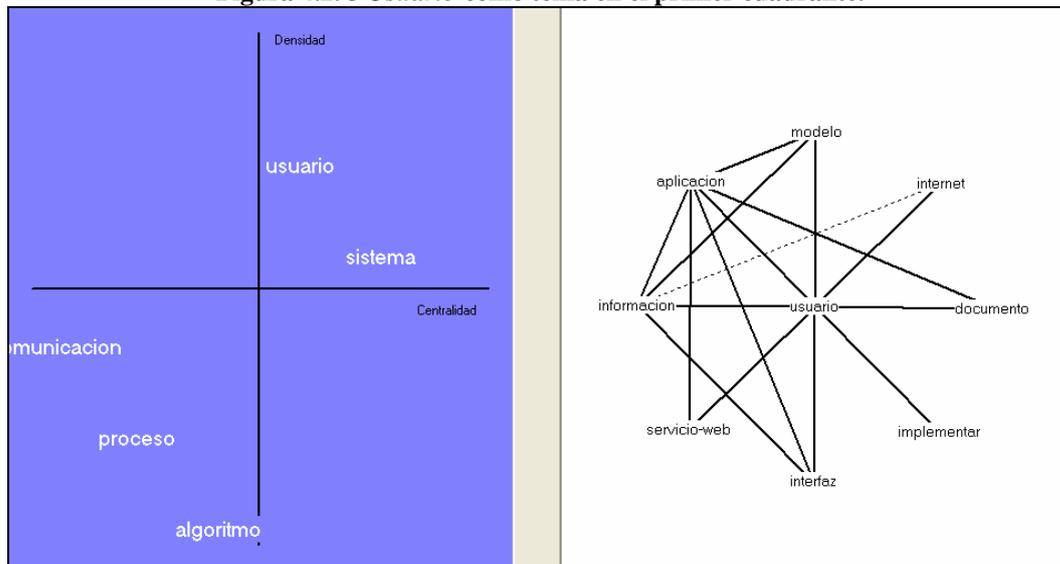
**Figura 4.2.2 Usuario como tema en el cuarto cuadrante.**



Fuente: Elaboración Propia con base de datos del CIC usando el software Redes 2005 v.1.0.0.2008

En la Figura 4.2.3, se muestra el tema *usuario* que corresponde al periodo 2005-2007, situándose en el primer cuadrante.

**Figura 4.2. 3 Usuario como tema en el primer cuadrante.**



**Fuente: Elaboración Propia con base de datos del CIC usando el software Redes 2005 v.1.0.0.2008**

Este análisis dinámico de un tópico de computación a través de varios periodos de tiempo, de los papeles (tema, descriptor) que toma a través de estos periodos y de la secuencia que siga entre los cuadrantes, indicará cuando un tópico se puede volver parte central de los desarrollos de tesis. Este comportamiento resultará de interés para poder predecir con mayor seguridad cuando un tópico se volverá parte central de los desarrollos de tesis.

## Conclusiones y Recomendaciones

Las conclusiones de esta tesis se separaran en dos clasificaciones, las propias relacionadas con la *cienciometría* y las correspondientes al área de aplicación.

Respecto a la *cienciometría*, las conclusiones son las siguientes:

1. El análisis *cienciométrico* permite al “aplicador” desarrollar en forma exhaustiva el análisis de los datos sobre la producción científica. Este trabajo de tesis permitió fundamentar conclusiones sobre el comportamiento de la producción científica, en este caso de temas de tesis de maestría. Conclusiones como el incremento de tesis dirigidas entre dos directores.
2. La *cienciometría* es una ciencia joven, con un desarrollo aproximado de 30 años y que se apoya de otra ciencia como la *bibliometría*, “joven” si la comparamos con cualquier otra ciencia como la química, las matemáticas o la física, algunas con cientos o miles de años, como cualquier ciencia en formación, constantemente se adicionan herramientas o técnicas de trabajo en su principal tarea de medir, que son herramientas de cómputo y las correspondientes técnicas matemáticas de medir y visualizar los resultados.
3. Para un análisis *cienciométrico*, se recomienda describir la fuente de datos con indicadores de publicación para conocer en detalle la actividad científica, también se recomienda realizar el análisis de citas para conocer las fuentes de conocimiento de la producción científica y localizar su posible “inspiración”.
4. Con el análisis dinámico de los tópicos de la computación, el área académica del centro puede identificar relaciones entre los tópicos que se estudian e investigan por la comunidad académica del CIC, además de revisar la actualidad de los tópicos incluidos en los desarrollos de tesis y para predecir o pronosticar cuando un tópico se vuelve parte central en el desarrollo de las tesis, lo que ayudará a satisfacer las necesidades informativas para orientar sus investigaciones, también en las empresas puede ayudar a generar alguna ventaja competitiva.

Y con respecto al área de aplicación se concluye lo siguiente:

1. El crecimiento en los últimos dos años de tesis codirigidas se puede ver debido a que el reglamento de posgrado del 2006 ya permite la codirección de tesis con dos investigadores del mismo centro, codirección que antes estaba condicionada a un investigador externo y a un investigador del centro. La mayoría de las codirecciones en los primeros años eran con profesores externos.
2. Solo tres laboratorios mantienen una colaboración con la mayoría de los laboratorios; la mayoría tiene colaboración con el grupo de investigadores externos, y cuatro laboratorios solo tienen colaboración con un laboratorio.
3. Los desarrollos de tesis están orientados a resolver problemas de usuarios, por los tópicos que sobresalen y toman fuerza y que se muestran como temas en los diagramas estratégicos, como: *sistema, programa, servidor, desarrollo, proceso, tecnología, modelo, usuario*, entre otros y muy pocos hablan sobre investigación.
4. Los alumnos del centro no citan correctamente en su gran mayoría los elementos mínimos necesarios que lleva una referencia como nombre del autor, título del documento, editorial, paginas, volumen, entre otros, por lo que al realizar el análisis de citas resulta ser una tarea muy desgastante. Debido a que no hay una estandarización, se propone seguir un estilo de citas como el APA ó el CHICAGO.
5. La antigüedad de las citas a libros, es relativamente reciente ya que casi el 70% tiene menos de doce años.

### **Líneas de Investigación a futuro**

- Como se indica en la Ley de Zipf, realizar una investigación para entender mejor los contenidos de los documentos como menciona (Luhn, 1958), es posible localizar la zona de las palabras donde se describa en forma semántica mejor un documento y así dar una explicación en cuanto a la riqueza del contenido del documento en el área de desarrollo.
- Otra investigación relacionada con formación de grupos de conocimiento (redes sociales), es mostrar la participación de grupos de trabajo en la dirección de tesis, para investigar las posibles causas de éxito en estas direcciones; todos son expertos del área, todos son doctores, varios directores tiene éxito en dirección continua, entre otras causas.
- Otra línea de investigación analizando la estructura de los documentos de tesis, es para definir los elementos mínimos necesarios y que debe de contener un documento de tesis, así como los criterios para calificar cada uno de estos elementos (claridad en el objetivo, formato de bibliografía correcto y completo, entre otros), para obtener una tesis de calidad.

---

## BIBLIOGRAFÍA

Carbonel, S. (2007). Producción de autores cubanos en las revistas sobre ciencias de la computación registradas en el *Journal Citation Report* en el periodo 1990-2005, *Acimed*, 15(5). Cuba.

Callon, M., Courtail, JP., Penan, H (1993). *Cienciometría*. Trea, España.

De Filippo, Fernandez, M. (n.d). *Bibliometría: Importancia de los Indicadores Bibliométricos*. Redes. Argentina.

Deleus, F., Van, HM (2003). Science and technology interactions discovered with a new topographic map-based visualization tool. Katholieke Universiteit Leuven Campus Gasthuisberg, Belgica.

Díaz, P. (1995). *Análisis de Redes Sociales aplicado a la producción científica forestal española*. Facultad de biblioteconomía y documentación. Universidad de Granada.

Ding, Y., Chowdhury, G., Foo, S. (2000). Organizing Keywords in a Web Search Environment: A Methodology Based On Co-word Analysis. Proc. 6<sup>th</sup> International Society for Knowledge Organization (ISKU 6) Conference, Toronto, Canada, 28-34.

Ding, Y., Chowdhury, G., Foo, S. (2001). Bibliometric Cartography of Information Retrieval Research by Using co-word Analysis. *Information Processing & Management*, Singapore, 37(817-842).

Escorcía, OT., Pout, PR., (2009). Análisis bibliométrico de los artículos originales publicados en la revista *Universitas Scientiarum* (1987-2007). *Universitas Scientiarum*. Bogotá.

Escorsa, P., Maspons, R (2001). De la Vigilancia Tecnológica a la Inteligencia Competitiva. Prentice Hall, *La Cienciometría y sus Aplicaciones* (87-104). Madrid.

Escorsa, P., Maspons, R (2001). De la Vigilancia Tecnológica a la Inteligencia Competitiva. Prentice Hall, *Los Mapas Tecnológicos* (105-130). Madrid.

Fernández, A., Torralbo, M., Rico, L., Gitiérrez, P., Maz, A. (2003) Análisis cienciométrico de las Tesis Doctorales Españolas en Educación Matemática. *Revista Española*, 26(2), 162-176.

Garfield, E., (2007). From the Science of Science to Scientometrics. Visualizing the History of Science with *HistCite* Software, 11th International Conference of the International Society for Scientometrics and Informetrics. *Scientometrics*, (1), 21-26.

Gorbea, PS., (2005). El modelo matemático de Lotka: Su aplicación a la producción científica latinoamericana en ciencias bibliotecológicas y de la información. México D.F, UNAM: Centro Universitario de Investigaciones Bibliotecológicas.

Jurado, AE., Bailón, MR., Ruiz, BR. (2002). Evaluación a través del análisis de palabras asociadas (I). Departamento de Ingeniería Química. Universidad de Granada.

---

Larreina, GS., Hernando, S., Grisaleña, D (2006). La evolución de la Inteligencia Competitiva: un estudio de las herramientas cuantitativas. *PUZZLE: Revista Hispana de la Inteligencia Competitiva*, 5(020), 5-9.

Luhn, H.P., (1958). The automatic creation of literature abstracts”, *IBM Journal of Research and Development*, New York. 159-165.

Moon, YH., (2004), *Monitoring and Early Warning of Technological Progress*, First Workshop of the Research Planning Assessment Study Society, KISTI.

Pineda, V., Manterola, C., Vial, M., Losada., H. (2005). ¿Cuál es la calidad metodológica de los artículos referentes a terapia publicados en la Revista Chilena de Cirugía? *Revista Chilena*. 57(6).

Ruiz, BR., (nd). Ley de Lotka. Departamento de biblioteconomía y documentación. Universidad de Granada. España.

Ruiz, BR., (nd). El Lenguaje y la ciencia. Ley de Zipf. Departamento de biblioteconomía y documentación. Universidad de Granada. España.

Sánchez, G., Calvet, C. (n.d). La bibliometría como una herramienta de la Bioinformática, Facultad de Ciencias. Universidad Nacional Autónoma de México (UNAM).

Sotolongo, G., Guzman, M., García, I., Sanz, E., (2001), *Retos de la bibliometría: La vigilancia y evaluación de la actividad científica-tecnológica.*, Instituto Finlay. Habana, Cuba.

Sung, S., Jung, M (2008), Knowledge sources of innovation studies in Korea: A citation analysis, *Scientometrics*, 75 (1), 01–18.

Spinak, E (1988). Indicadores Cuantitativos. Seminario sobre evaluación de la producción científica. Brasil.

Vinkler, P (1988). An Attempt of surveying and classifying bibliometric indicator for scientometric purposes. *Scientometrics*, 13(5-6), 239-259.

Wyoung, WL (2007). How to identify emerging research fields using scientometrics: An example in the field of Information Security. *Scientometrics*, 76(3), 503-525

Zacca, GG., (2010) *Producción científica latinoamericana en la Salud Pública*. Trabajo de Grado, Diploma de Estudios Avanzados, Universidad de Granada y Universidad de la Habana. Habana, Cuba.

Zhu, D., Porter, A (1999). A process for mining science & technology documents databases, illustrated for the case of knowledge discovery and data mining, *Technology Policy & Assessment Center*, Georgia Institute of Technology, Atlanta, GA. 30332-0205

## **CIBERTECA**

<http://www.histcite.com/>

Consultada el 28 de agosto de 2008

---

<http://www.histcite.com/whatcanitdo.htm>  
Consultada el 06 de octubre de 2008

[http://www.zaintek.net/ebizkaia/Cursos/Curso\\_1/menu/pdf/3.pdf](http://www.zaintek.net/ebizkaia/Cursos/Curso_1/menu/pdf/3.pdf)  
Consultada el 10 de Julio 2009

<http://www.ugr.es/~rruizb/cognosfera/index.htm>  
Journal Of Citation Report: Consultada el 28 de agosto de 2008  
<http://www.uned.es/biblioteca/basesdedatos/jcr.pdf>

Larreina, S., Hernando, S., Griselda, D. (1995). La Evolución de la inteligencia competitiva: un estudio de las herramientas cuantitativas. Revistas Puzzle. Consultada el 25 de agosto de 2008, <http://www.leia.es/esp/publicaciones/sexta.htm>

<http://medios1.tripod.com/mediospe/id15.html>

<http://spanish.martinvarsavsky.net/general/producto-bruto.html>

<http://informationr.net/ir/9-4/paper193.html>

<http://www.bibliociencias.cu/gsd/collect/eventos/index/assoc/HASH013f/26f1bc40.dir/doc.pdf>

[http://es.wikipedia.org/wiki/Ley\\_de\\_Zipf](http://es.wikipedia.org/wiki/Ley_de_Zipf)

[http://www.ugr.es/~rruizb/cognosfera/sala\\_de\\_estudio/bibliometria/lotka.ppt](http://www.ugr.es/~rruizb/cognosfera/sala_de_estudio/bibliometria/lotka.ppt)

[http://www.ugr.es/~rruizb/cognosfera/sala\\_de\\_estudio/bibliometria/zipf.ppt](http://www.ugr.es/~rruizb/cognosfera/sala_de_estudio/bibliometria/zipf.ppt)

## ANEXO 1. BASE DE DATOS EN LOS TEMAS DE TESIS

Con los datos que se muestran en las siguientes tres tablas se procede a realizar los análisis para obtener los indicadores de publicación y de citación.

### A1.1 Tesis Desarrolladas durante el periodo 1996-2007

No	NOMBRE DEL ALUMNO	FECHA EXAMEN	TEMA DE TESIS	DIRECTOR	CODIRECTOR
1	García Mayer Mario Rolando	09/05/1997	Sistema reporteador para info	Sandoval Reyes Sergio	
2	Monroy Santamaría Edgar	10/10/1997	Sistema reporteador automático	Galindo Soria Leopoldo Alberto	
3	Germán Ibarra Alberto Alfonso	12/12/1997	Sistema computarizado de au	Galindo Soria Leopoldo Alberto	
4	Hernández Gutiérrez Carlos	03/11/1998	Sistema gráfico para edición	Sandoval Reyes Sergio	
5	Chac Kantún Martín Guadalupe	17/03/1999	Herramientas para el agrupar	Ruiz Shulcloper José	
6	González García Eduardo	19/03/1999	Herramientas para el proceso	Maldonado Salazar Tomás Juan Bosco	
7	Palomino Ramírez María Cecilia	12/04/1999	Sistema general de gestión	el Guzmán Arenas Adolfo	
8	Paredes Vargas Carlos Francisco	16/04/1999	Manipulador topográfico de la	Sosa Iglesias José de Jesús Emilio	
9	Taddei Zavala Edmundo	11/05/1999	Un algoritmo genético para la	Flores Pérez Pedro	
10	Martínez Casas Marcos Gerardo	12/05/1999	Extensión basada en patrone	Ferro Castro Bárbaro Jorge	
11	Mendoza Austria Luis	21/05/1999	Herramientas de graficación	y Menchaca García Felipe Rolando	
12	Carbajal Rivera Ernesto	31/05/1999	Servicio de autenticación en r	Coyote Estrada Hugo César	
13	Rivera Santos Ivonne	11/06/1999	TERTRIES: Una nueva estruc	Coyote Estrada Hugo César	
14	Avila Rosas Arturo	16/06/1999	Desarrollo de un servidor con	Sosa Iglesias Jesús Emilio	Ferro Castro Bárbaro
15	Rodríguez Ramírez José Luis	23/06/1999	Modelado tridimensional de e	Cen Zubieta José Rafael	
16	León Enríquez Javier	25/06/1999	Sistema de adquisición de da	Suárez Guerra Sergio	Isaac Schifter Secora
17	Durán Martínez Jorge Pedro	06/07/1999	Sistema de monitoreo y adqu	Suárez Guerra Sergio	
18	Rodríguez García José Luis	06/08/1999	Modelo de trabajo grupal y ev	Núñez Esquer Gustavo	
19	Téllez Castillo Germán	20/08/1999	Los fractales, un punto de cor	González Hernández Manuel	Guzmán Rentería Ces
20	Curbelo García Ariadna	26/08/1999	Sistema para el control de la i	Ferro Castro Bárbaro Jorge	
21	Menchaca Méndez Rolando	27/08/1999	Un ORB para ADORE.	Coyote Estrada Hugo César	
22	Moncada Bolón Juan Jesús	10/09/1999	Extensión del concepto de im	Díaz de León Santiago Juan Luis	
23	Silva López Héctor Eduardo	26/10/1999	Sistema de adquisición y proc	Ferro Castro Bárbaro Jorge	Calderón Hernández I
24	Botello Castillo Alejandro	28/10/1999	Controlador ODBC para el G	Coyote Estrada Hugo César	
25	Pedro Guevara López	29/10/1999	Control de motores de corrier	Suárez Guerra Sergio	Medel Juárez José de
26	Christos Lazaridis Kourtis Chrysant	29/11/1999	Analizador estadístico de text	Mikhail Alexandrov Aronovich	
27	Orantes Jiménez Sandra Dinora	06/12/1999	Patrones de diseño para inter	Ferro Castro Bárbaro Jorge	
28	Quintero Téllez Rolando	07/01/2000	Asignación de comportamiento	Alvarado Mentado José Matías	Núñez Esquer Gustav
29	Polo Olivares José Antonio	14/02/2000	Sistema creador de estructura	Galindo Soria Leopoldo Alberto	Velazquez Aurelio
30	Hernández Cruz Venancio Eric	09/03/2000	Desarrollo de un sistema com	Menchaca García Felipe Rolando	
31	Escamilla Velázquez Gelmunt Lorer	15/03/2000	Aplicación de un sistema exp	González Hernández Manuel	
32	Villalobos Martínez Angélica	18/05/2000	Simulación de un sistema inte	Sandoval Reyes Sergio	Alfaro Pastor Ernesto
33	Manjarrez Sánchez Jorge Roberto	23/05/2000	Controlador JDBC tipo 3 para	Ferro Castro Bárbaro Jorge	
34	Espinosa Cruz Andrés Avelino	16/06/2000	Sistema de autenticación y c	Coyote Estrada Hugo César	
35	Balladares Ocaña Leandro	27/06/2000	Ambiente CASE y lenguaje vi	Núñez Esquer Gustavo	Alvarado Mentado Jos
36	Eslava García Georgina	22/08/2000	Análisis de fractalidad de agr	Lira Galeana Carlos	Coyote Estrada Hugo
37	Sotelo Rangel Rafael	03/10/2000	Metodología para el conteo d	Sossa Azuela Juan Humberto	
38	Magadán Méndez Margarita Julia	23/10/2000	Errores de definición en los m	Díaz de León Santiago Juan Luis	
39	Armenta Alcántara Andrés Antonio	30/10/2000	Asignación de comportamiento	Núñez Esquer Gustavo	
40	Castillo Santana Rutilio Ricardo	24/11/2000	Sistema de costos de mano d	Sosa Iglesias Jesús Emilio	Cen Zubieta José Raf
41	Leyva del Follo Luis Eduardo	11/12/2000	Análisis de algoritmos y arqu	Menchaca García Felipe Rolando	
42	Ríos Suriano Francisco Javier	14/12/2000	Estabilización dinámica de sis	López Morales Virgilio	
43	Piñón Méndez Guadalupe	14/12/2000	Optimización de la fragmenta	Kuri Morales Ángel Fernando	
44	Reyes Villagrán Armando Javier	15/12/2000	La videoconferencia de escrit	Román Nájera Susana Mónica	
45	Herrera Alcántara Oscar	15/12/2000	Compresión y descompresión	Ferro Castro Bárbaro Jorge	
46	Cortés Galicia Jorge	05/03/2001	Sistema e protocolos con apli	Menchaca García Felipe Rolando	
47	Gama Moreno Luis Antonio	05/03/2001	Sistema de punto de venta pa	Menchaca García Felipe Rolando	
48	García Casas Samuel	04/04/2001	Sistema computacional para	Galindo Soria Leopoldo Alberto	
49	Amorós Zavala Evaluz Citlali	06/04/2001	Módulo de evaluación multicri	Ramos Rodríguez José María	
50	Nango Solís Galdino Belizario	26/04/2001	Desarrollo de un intérprete pa	López Morales Virgilio	

**Fuente:** Elaboración Propia con base de datos del CIC .2008

## Continuación de la Tabla A.1.1

51	Escudero Miranda Luis Alberto	05/05/2001	Ambiente distribuido para ope	Menchaca García Felipe Rolando
52	Torres Ruiz Miguel Jesús	07/05/2001	Herramienta SIG de escritoric	Ramos Rodríguez José María
53	Vicuña Heredia Juan Roberto	18/05/2001	Guía para el desarrollo del ad	Guevara López Pedro
54	Pérez Araiza Gonzalo	07/06/2001	Control de acceso y de prese	Menchaca García Felipe Rolando
55	Casillas Pérez Gilberto Armando	29/08/2001	Automatización de cálculos p	Blanco Cano Xóchitl Guillermina Sandoval Reyes Serg
56	González Palacios Jorge Luis	31/08/2001	Desarrollo y construcción de	Cheremetov Leonid
57	Moreno Ibarra Marco Antonio	04/09/2001	La generalización automática	Viniegra Islas Rodolfo Alfredo
58	González Santos Leopoldo	07/09/2001	Segmentación semiautomátic	Barrios Álvarez Fernando Alejandr
59	Vargas de Basterra Ricardo	28/09/2001	Análisis de información en ár	Ferro Castro Bárbaro Jorge
60	Vázquez Romero Sandra	05/10/2001	Estudio de factibilidad técnica	Cheremetov Leonid
61	González Sánchez Arturo	19/10/2001	Aplicación de redes neuronale	Albornoz Bueno Álvaro José
62	Silva Fernández Jesús Martín	23/10/2001	Modelo agente de recuperació	Cheremetov Leonid
63	Moncayo López Luis Miguel	09/11/2001	Ambiente de diseño de redes	Menchaca García Felipe Rolando
64	López Paredes Enrique	21/11/2001	Seguimiento de objetos usan	Sossa Azuela Juan Humberto
65	Xicoténcatl Aguilar Karla Penélope	21/11/2001	Conteo de objetos en escena	Sossa Azuela Juan Humberto
66	Perales Briseño Alejandro	26/11/2001	Diseño y desarrollo de un osc	Barrón Fernández Ricardo
67	Dorantes Huerta Daniel	26/11/2001	Redes neuronales para el mo	Kuri Morales Ángel Fernando
68	Decelis Burguete Joaquín Octavio	29/11/2001	Digitalización automatizada d	Levachkine Serguei Pavlovich
69	Sánchez Rodríguez Narda Inés	30/11/2001	Sistema para el análisis y dis	Menchaca García Felipe Rolando
70	Ávila Martínez José Antonio	05/12/2001	Desarrollo de fronts-ends en	Sosa Iglesias José de Jesús Emilio Ferro Castro Bárbaro
71	Morales Pacheco Roberto Antonio	05/12/2001	Simulación de agregados asf	Virgilio López Morales
72	Domínguez Domínguez Janette	07/12/2001	Integración no supervisada d	Kuri Morales Ángel Fernando
73	Olguín Salguero Juan Miguel	10/12/2001	Sistema de graficación de obj	Ferro Castro Bárbaro Jorge
74	Santaolalla Salgado Javier	17/12/2001	Concepción de un modelo pa	Gutiérrez Tornés Agustín Francisco
75	Contreras Montoya Miguel	11/04/2002	Desarrollo de una infraestruct	Cheremetov Leonid
76	Rico Jiménez Blanca Alicia	24/05/2002	Simulación en tiempo real de	Guevara López Pedro
77	Chanona Hernández Liliana	28/05/2002	Compilación de un corpus re	Guelboukh Kahn Alexandre Felixovich
78	Velázquez Cabrera Roberto	07/06/2002	Estudio de aerófonos mexica	Cen Zubieta José Rafael
79	Trejo Soto Gloria Irene	14/06/2002	Módulo XML para acceder al	Orantes Jiménez Sandra Dinora
80	Domínguez Ayala María del Carme	27/06/2002	Manejo de infinitos eventos in	Guzmán Arenas Adolfo
81	Henestrosa Carrasco Leticia	28/06/2002	Desarrollo de la metodología	Cheremetov Leonid
82	Larios Osorio Martín	31/07/2002	Bodega de datos con interfac	Sandoval Reyes Sergio
83	Jurado Jiménez Roberto	24/09/2002	La aplicación de métodos de	Mikhail Alexandrov Aronovich Guelboukh Khan Alex
84	Vázquez Parada Ilyana Verónica	30/10/2002	Sistema generador de datos r	Martínez Luna Gilberto Lorenzo
85	Ramos García Mario Eduardo	13/12/2002	Mejoramiento de imágenes di	Felipe Riverón Edgardo Manuel
86	García Robles Liseth	13/12/2002	Sistema para la gestión y aná	Suárez Guerra Sergio
87	Ayaquica Martínez Irene Olaya	16/12/2002	Algoritmo C-MEANS difuso us	Martínez Trinidad José Francisco Velazquez Aurelio
88	Reyes Damián Lucas	17/12/2002	Visualización y control remot	Ferro Castro Bárbaro Jorge
89	Germán Soto Ernesto	17/12/2002	Desarrollo de una herramient	Cheremetov Leonid
90	Sánchez Herrera Jorge	10/01/2003	Búsqueda tabú para resolver	Morales Mendoza Luis Bernardo Aguilar Ibañez Carlos
91	Pruneda Martínez Arturo	13/02/2003	Desarrollo de una infraestruct	Menchaca Méndez Rolando
92	Dávalos Rodríguez Felipe	13/03/2003	Clasificador de documentos l	Sosa Iglesias José de Jesús Emilio Cen Zubieta José Raf
93	Guzmán Lugo José Giovanni	14/03/2003	Metodología extendida para e	Sossa Azuela Juan Humberto
94	Romero Alvario Adolfo	04/04/2003	Herramienta para la obtenció	Ferro Castro Bárbaro Jorge
95	Varela García Francisco	07/04/2003	Diagnosticador de fallas basa	Sandoval Reyes Sergio
96	Langurén Rodríguez Alejandro	25/04/2003	Modelo de lealtad.	Ferro Castro Bárbaro Jorge
97	Rivera Loza Gabriela	04/06/2003	Selección automática e primit	Guelboukh Khan Alexandre Felixovich
98	De la Cruz Sánchez Héctor	30/06/2003	Biblioteca de agentes para la	Gutiérrez Tornés Agustín Francisc Tirado Rios Uriel
99	Miranda Jiménez Sabino	01/07/2003	El formateo de entrada y la c	Bolshakov Miranova Igor
100	Castillo Velásquez Francisco Antoni	01/07/2003	Sistema de análisis morfológi	Sidorov Grigori

## Continuación de la Tabla A.1.1

101	Cuevas Rasgado Alma Delia	03/07/2003	La evaluación de los productos	Gutiérrez Tornés Agustín Francisco
102	Sarmiento Tovilla Juan Carlos	04/07/2003	Un modelo del sistema inmune	Ferro Castro Bárbaro Jorge
103	García Márquez Marilupe	28/08/2003	Ensayo del modelo de calidad	Martínez Luna Gilberto Lorenzo
104	Quiane Ruiz Jorge Arnulfo	29/08/2003	Diseño e Implementación de un	Manjarrez Sánchez Jorge Roberto
105	Álvarez Robles Laura	21/10/2003	Sistema de asignación de recursos	Menchaca Méndez Rolando
106	Santiago Montero Raúl	24/10/2003	Clasificador Híbrido de patrones	Yáñez Márquez Cornelio
107	Calderón Sambarino María Julia	13/11/2003	Sistema para el estudio, medición	Suárez Guerra Sergio
108	Naranjo Ramírez Gerardo	15/12/2003	Herramienta computacional e	Galindo Soria Leopoldo Alberto
109	Aragón Pineda Enrique de Jesús	15/12/2003	Controlador JBDC para múltiples	May Arrijoa Alfonso de los Ángeles
110	Juárez Rodríguez Felipe	18/12/2003	Sistema experto basado en reglas	Ferro Castro Bárbaro Jorge
111	Argüelles Centeno Salvador Enrique	18/12/2003	Sistema para la regionalización	Sossa Azuela Juan Humberto
112	Carreto Arellano Chadwick	16/02/2004	Modelo y arquitectura para la	Menchaca Méndez Rolando
113	Couder Castañeda Carlos	05/03/2004	Simulación de un flujo supersónico	Ferro Castro Bárbaro Jorge
114	Paredes Aguirre Daniel Eduardo	08/03/2004	Servicios de localización sobre	Menchaca Méndez Rolando
115	Valladares Solís Enrique Ignacio	12/03/2004	Sistema de evaluación en línea	Ferro Castro Bárbaro Jorge
116	Ramos Ramos Rafael	03/05/2004	Diseño e implementación de un	Barrón Fernández Ricardo
117	Benítez Benítez Héctor	14/05/2004	Cálculo de momentos geométricos	Sossa Azuela Juan Humberto
118	González Gómez Efrén	02/07/2004	Reconocimiento de patrones	Levachkine Serguei Pavlovich
119	Lara Gaytán Leonardo	19/08/2004	Metodología de dimensionamiento	Menchaca García Felipe Rolando
120	Hernández Rubio Erika	19/08/2004	Compilación automática del código	Sidorov Grigori
121	Cruz Pérez Daniel	27/08/2004	Modelo dinámico para tareas	Guevara López Pedro
122	Rivera Blas Emmanuel Zenén	14/10/2004	Conversión de archivos con	Cen Zubieta José Rafael
123	Cortés Ruiz Felisa	05/11/2004	Diseño e implementación de un	Gutiérrez Tornés Agustín Francisco
124	Medina Meneses Víctor Manuel	22/11/2004	Instrumento Virtual para la	Suárez Guerra Sergio
125	Morales Gutiérrez Sandra Luz	13/12/2004	Detección automática de la	Felipe Riverón Edgardo Manuel
126	Jiménez Contreras Edith Adriana	13/12/2004	Modelo para estimación de	Ferro Castro Bárbaro Jorge
127	Sánchez Garfias Flavio Arturo	15/12/2004	Lernmatrix de Steinbuch: Con	Yáñez Márquez Cornelio
128	Jiménez Hernández Hugo	17/12/2004	Reconocimiento de caracteres	Figueroa Nazuno Jesús Guillermo
129	Flores Mejía Eduardo	17/12/2004	Sistema de instalación de software	Quintero Téllez Rolando
130	Castillo Díaz Víctor Daniel	17/12/2004	μCronos: Un micronúcleo para	Menchaca Méndez Rolando
131	Sánchez Cerón Marco Antonio	19/01/2005	Sistemas de Información para	Sandoval Reyes Sergio
132	Medina Apodaca Juan Manuel	15/02/2005	Herramienta para la extracción	Figueroa Nazuno Jesús Guillermo
133	Mata Rivera Félix	25/02/2005	Recuperando y analizando	Torres Ruiz Miguel Jesús
134	Menchaca Méndez Ricardo	01/03/2005	Una herramienta para el estudio	Figueroa Nazuno Jesús Guillermo
135	Avila Sánchez Christian Alejandro	17/03/2005	Extracción automática de	Figueroa Nazuno Jesús Guillermo
136	González Escalante Luis	04/05/2005	Interfaz para la publicación de	Botello Castillo Alejandro
137	Guerra Vázquez José Juan	17/06/2005	Un sistema multiagente para	Cheremetov Leonid
138	Salgado Ramírez Julio César	20/06/2005	Estudio estadístico comparativo	Yáñez Márquez Cornelio
139	López Ruiz Gabriela de Jesús	05/07/2005	Análisis de algoritmos de	Medel Juárez José de Jesús
140	Rojas Hernández Rafael	15/07/2005	Compresión de imágenes con	Díaz de León Santiago Juan Luis
141	Ortiz Montes Alicia	31/10/2005	Momentos geométricos de	Díaz de León Santiago Juan Luis
142	Olascoaga Vergara Blanca Rocío	17/11/2005	Sistema de pruebas de	Gutiérrez Tornés Agustín Francisco
143	Cruz Torres Benjamín	28/11/2005	Restauración de palabras	Sossa Azuela Juan Humberto
144	Vera Félix José Ángel	16/12/2005	Sistema de detección	Sidorov Grigori
145	Castro Sánchez Noé Alejandro	16/12/2005	Sistema de análisis automático	Sidorov Grigori
146	Vázquez Espinoza de los Monteros	19/12/2005	Reconocimiento de objetos	Sossa Azuela Juan Humberto
147	Carvajal Hernández José Juan	17/01/2006	Sistema distribuido para	Sánchez Fernández Luis Pastor
148	Castro Medina Gabriel	03/02/2006	Metodología integral para el	Orantes Jiménez Sandra Dinora
149	Cortés León Héctor	10/02/2006	Relevador digital de	Barrón Fernández Ricardo
150	Escamilla Javier Cuauhtle	17/03/2006	Desarrollo de un sistema de	Suárez Guerra Sergio

## Continuación de la Tabla A.1.1

151	Alvarado Cervantes Rodolfo	30/03/2006	Segmentación de patrones lin	Levachkine Serguei Pavlovich	
152	Pineda Briceño Anabel	17/04/2006	Arquitectura de control ubicuc	Menchaca Méndez Rolando	Guzmán Lugo José G
153	Rodríguez Lizalde José Marcelo	22/05/2006	Clasificación de series de tien	Figuroa Nazuno Jesús Guillermo	
154	Cuevas Cabrera Walter	26/05/2006	Módulo de soporte a la toma	Orantes Jiménez Sandra Dinora	Botello Castillo Alejan
155	Isaí Fararoni Rodríguez	26/05/2006	Teleoperación de robots CNC	Díaz de León Santiago Juan Luis	
156	Maya Cilia Juan	01/06/2006	Sincronización de semáforos	Olivares Ceja Jesús Manuel	
157	Landassuri Moreno Víctor Manuel	06/06/2006	Predicción de series de tiemp	Figuroa Nazuno Jesús Guillermo	
158	Martínez Rosales Miguel Alejandro	07/06/2006	Descriptor topológico para m	Moreno Ibarra Marco Antonio	
159	Solís Estrella Habacuc	07/06/2006	Estudio comparativo de técnic	Gutiérrez Aldana Alfonso	
160	Tejada Cárcamo Javier Leandro	08/06/2006	Desambiguación de sentidos	Guelboukh Kahn Alexandré Felixovich	
161	De la Cruz Santiago Juan Carlos	08/06/2006	Orquestador de servicios WE	Botello Castillo Alejandro	
162	López Hinojosa Marisol	14/06/2006	Un nuevo algoritmo en la técr	González Santos José Germán	Guzmán Arenas Adol
163	Torres Ramos Sulema	22/06/2006	Aprendizaje supervisado de c	Guelboukh Kahn Alexandré Felixovich	
164	Olivas Zazueta Omar Alejandro	23/06/2006	Sistema de construcción de r	Sidorov Grigori	
165	Gallegos Gallegos Patricia	28/06/2006	Diseño del sistema de comun	Gutiérrez Aldana Alfonso	
166	Martínez Valle Carlos León	29/06/2006	Elaboración de un modelo de	Orantes Jiménez Sandra Dinora	
167	Ángeles Yreta Mario Alberto	30/06/2006	Cómputo de la similitud entre	Figuroa Nazuno Jesús Guillermo	
168	Montiel Hernández Noé Maxihelleer	01/08/2006	Posicionamiento basado en r	Moreno Ibarra Marco Antonio	
169	Luna Benoso Benjamín	25/09/2006	Análisis y generalización de u	Téllez Castillo Germán	
170	Flores Carapia Rolando	27/11/2006	Memorias asociativas alfa-be	Yáñez Márquez Cornelio	
171	Gutiérrez Álvarez Telésforo Emilian	19/12/2006	Ambiente ubicuo para colaboi	Menchaca Méndez Rolando	Quintero Téllez Rolan
172	Fabela Cano Oscar	17/01/2007	Biblioteca digital basada en s	Peredo Valderrama Rubén	
173	Rodríguez Robledo Gricelda	16/03/2007	Sistema ubicuo de historia clí	Olivares Ceja Jesús Manuel	
174	Cruz Meza María Elena	23/03/2007	Aprendizaje y recuperación d	Yáñez Márquez Cornelio	Garfías Sánchez Flav
175	Torrescano Arias Daniel Aarón	20/04/2007	Estándar de hardware y softw	Díaz de León Santiago Juan Luis	
176	Silva Vidal Rodiver	27/04/2007	Sistema de administración de	Menchaca Méndez Rolando	
177	Gómez Balderas José Ernesto	22/05/2007	Resolución automática de la	Sidorov Grigori	Jimenez Salazar Héct
178	Horta Mendoza Juan Manuel	04/06/2007	Ambiente visual de simulació	Guzmán Lugo José Giovanni	Menchaca Méndez R
179	Lugo García Tania	07/06/2007	Aprendizaje automático de la	Guelboukh Kahn Alexandre Felixovich	
180	Garro Licón Beatriz Aurora	13/06/2007	Estudio comparativo de difere	Sossa Azuela Juan Humberto	
181	Hernández Atilano Roberto	15/06/2007	Identificación de vehículo em	Sandoval reyes Sergio	
182	Verástegui Barranco Karina	21/06/2007	Conceptualización de aspectc	Moreno Ibarra Marco Antonio	Torres Tuíz Miguel Je
183	Novoa Cataño Javier	22/06/2007	Análisis de la probabilidad de	Téllez Castillo Germán	
184	Catalán Salgado Edgar Armando	29/06/2007	Memorias asociativas alfa-be	Yáñez Márquez Cornelio	
185	Román Godínez Israel	03/07/2007	Aplicación de los modelos apl	Yáñez Márquez Cornelio	
186	González Arroyo Lilia	04/07/2007	Componentes multimedia orie	Peredo Valderrama Rubén	
187	Martínez Díaz Arturo Benjamín	04/07/2007	Diseño de un modelo espacio	Torres Ruiz Miguel Jesús	Moreno Ibarra Marco
188	Villegas Juárez Virginia	04/07/2007	Identificación de conceptos g	Quintero Téllez Rolando	
189	Hernández Contreras Euler	04/07/2007	Sistema multiagente para el	a Cheremetov Leonid	Oleksiy Pogrebnyak
190	López Yáñez Itzamá	05/07/2007	Clasificador automático de alt	Figuroa Nazuno Jesús Guillermo	Yáñez Marquéz Corn
191	Montes de Oca Morales Víctor	05/07/2007	Plantilla para la representació	Torres Ruiz Miguel Jesús	Moreno Ibarra Marco
192	Olgún Aguilar Luis Antonio	06/07/2007	Arañas distribuidas: asignació	Guzmán Arenas Adolfo	
193	Peredo Valderrama Iván	30/07/2007	Sistema cliente/servidor para	Balladares Ocaña Leandro	
194	Casarrubias Vargas Heriberto	31/07/2007	Generación de trayectorias p	e Sossa Azuela Juan Humberto	Barrón Fernández Ric
195	Petrilli Barceló Alberto Elías	31/07/2007	Control visual de un robot mó	Sossa Azuela Juan Humberto	Barrón Fernández Ric
196	Ortega González Eric Vladimir	09/08/2007	Una técnica para el análisis d	Figuroa Nazuno Jesús Guillermo	
197	Ramírez Amaro Karinne	23/11/2007	Técnicas de aprendizaje de e	Chimal Eguía Juan Carlos	
198	Hernández Cruz Macario	05/12/2007	Generador de los grafos conc	Guelboukh Kahn Alexandre Felixovich	
199	Martínez Muñoz Jorge	01/02/2008	Sistema multi-agente para m	c Alvarado mentado José Matías	
200	Estrada Segovia Guadalupe Manue	29/02/2008	Arquitectura para la construc	Quintero Téllez Rolando	
201	Mirón Bernal Miguel Ángel	06/03/2008	Análisis de la similitud entre	p Coyote Estrada Hugo César	Figuroa Nazuno Jes
202	Selley Rojas Héctor Julián	22/05/2008	Monitoreo del comportamiento	Menchaca García Felipe Rolando	

Tabla A1.2 Laboratorio de directores de Tesis

No	DIRECTOR	LABORATORIO 1
1	González Hernández Manuel	CINVESTAV
2	Alvarado mentado José Matías	LIA
3	Menchaca Méndez Rolando	LRyT
4	Martínez Trinidad José Francisco	INAOE
5	Guzmán Arenas Adolfo	LBDyTS
70	Viniegra Islas Rodolfo Alfredo	LG
71	Yáñez Márquez Cornelio	LI

Fuente: Elaboración Propia con base de datos del CIC.2008

Tabla A1.3 Tesis con dos directores

No	DIRECTOR	CODIRECTOR
1	Morales Mendoza Luis Bernardo	Aguilar Ibañez Carlos
2	Sandoval Reyes Sergio	Alfaro Pastor Ernesto
3	Núñez Esquer Gustavo	Alvarado Mentado José Matías
4	Cheremetov Leonid	Balladares Ocaña Leandro
5	Sossa Azuela Juan Humberto	Barrón Fernández Ricardo
6	Sossa Azuela Juan Humberto	Barrón Fernández Ricardo
7	Sidorov Grigori	Bolshakov Miranova Igor
8	Orantes Jiménez Sandra Dinora	Botello Castillo Alejandro
9	Ferro Castro Bárbaro Jorge	Calderón Hernández Isaac
10	Sosa Iglesias Jesús Emilio	Cen Zubieta José Rafael
11	Sosa Iglesias José de Jesús Emilio	Cen Zubieta José Rafael
12	Lira Galeana Carlos	Coyote Estrada Hugo César
13	Ferro Castro Bárbaro Jorge	De Albornoz Bueno Alvaro
14	Yáñez Márquez Cornelio	Díaz de León Santiago Juan Luis
15	Yáñez Márquez Cornelio	Díaz de León Santiago Juan Luis
16	Yáñez Márquez Cornelio	Díaz de León Santiago Juan Luis
17	Sosa Iglesias Jesús Emilio	Ferro Castro Bárbaro Jorge
18	Sosa Iglesias José de Jesús Emilio	Ferro Castro Bárbaro Jorge
19	Yáñez Márquez Cornelio	Garfías Sánchez Flavio Arturo
20	Mikhail Alexandrov Aronovich	Guelboukh Khan Alexandre Felixovitch
21	Medel Juárez José de Jesús	Guevara López Pedro
22	Ferro Castro Bárbaro Jorge	Gutiérrez Tornés Agustín Francisco
23	May Arriola Alfonso de los Angeles	Gutiérrez Tornés Agustín Francisco
24	Ferro Castro Bárbaro Jorge	Gutiérrez Tornés Agustín Francisco
25	Ferro Castro Bárbaro Jorge	Gutiérrez Tornés Agustín Francisco
26	González Santos José Germán	Guzmán Arenas Adolfo
27	Menchaca Méndez Rolando	Guzmán Lugo José Giovanni
28	González Hernández Manuel	Guzmán Renteria Cesar Saúl
29	Suárez Guerra Sergio	Isaac Schifter Secora
30	Sidorov Grigori	Jimenez Salazar Héctor
31	Suárez Guerra Sergio	Medel Juárez José de Jesús
32	Guevara López Pedro	Medel Juárez José de Jesús
33	Guzmán Lugo José Giovanni	Menchaca Méndez Rolando
34	Torres Ruiz Miguel Jesús	Moreno Ibarra Marco Antonio
35	Torres Ruiz Miguel Jesús	Moreno Ibarra Marco Antonio
36	Alvarado Mentado José Matías	Núñez Esquer Gustavo
37	Cheremetov Leonid	Oleksiy Pogrebnyak
38	Menchaca Méndez Rolando	Quintero Téllez Rolando
39	Blanco Cano Xóchitl Guillermina	Sandoval Reyes Sergio
40	Gutiérrez Tornés Agustín Francisco	Santaolaya Salgado René
41	Barrios Álvarez Fernando Alejandro	Sossa Azuela Juan Humberto
42	Gutiérrez Tornés Agustín Francisco	Tirado Rios Uriel
43	Moreno Ibarra Marco Antonio	Torres Tuíz Miguel Jesús
44	Galindo Soria Leopoldo Alberto	Velazquez Aurelio
45	Martínez Trinidad José Francisco	Velazquez Aurelio
46	Díaz de León Santiago Juan Luis	Yáñez Marquéz Cornelio
47	Díaz de León Santiago Juan Luis	Yáñez Marquéz Cornelio
48	Figuerola Nazuno Jesús Guillermo	Yáñez Marquéz Cornelio

Fuente: Elaboración Propia con base de datos del CIC.2008

**ANEXO 2. ESTADO Y TRATAMIENTO DE CITAS EN LOS DOCUMENTOS DE TESIS DE MAESTRÍA**

En los documentos de tesis se tienen cinco tipos de recursos de información:

- a) Libros
- b) Artículos
- c) Conferencias
- d) Páginas Electrónicas.
- e) otros (tesis)

Una primera revisión al número de citas en los documentos mostraron los resultados que se pueden ver en la Tabla A.1.1

**Tabla A.1.1 Citas por Documentos**

Tesis	Libros	Artículos	Conferencias	Electrónicas	otros (tesis)	Total
1	0	239	0	0	0	239
2	3	0	0	12	3	18
3	5	4	0	4	6	19
4	5	6	0	2	0	13
5	5	0	0	11	2	18
6	6	25	0	1	1	33
7	7	13	0	0	1	21
8	8	6	0	1	3	18
9	8	11	0	1	5	25
10	8	16	0	8	1	33
11	8	6	0	5	3	22
12	8	4	0	37	0	49
14	9	11	0	3	5	28
15	9	1	0	26	0	36
16	10	0	0	41	0	51
17	10	33	0	5	0	48
18	11	3	1	3	0	18
19	12	19	0	3	9	43
20	16	25	0	18	0	59
21	18	0	0	32	1	51
22	18	28	0	0	3	49
23	19	2	0	0	8	29
24	19	5	0	4	0	28
25	20	0	0	12	0	32
27	24	4	1	12	1	42
28	25	38	0	0	0	63
29	28	7	0	10	6	51
30	30	0	0	0	2	32
31	50	12	0	38	0	100
32	57	17	0	51	0	125
	<b>456</b>	<b>535</b>	<b>2</b>	<b>340</b>	<b>60</b>	<b>1393</b>

Fuente: Elaboración Propia con base de datos del CIC.2008

Las Figuras A.1.1, A.1.2, A.1.3, A.1.4 y A.1.5 muestran ejemplos del estado de cómo son mal citados los documentos de la Tabla A.1.1, donde se detecta el mal formato de una cita a un artículo, errores en la escritura de autores, falta de datos en la cita, entre otros errores que dificultan el almacenar los datos en la herramienta HistCite versión 8.5.26, y donde se requiere enfatizar sobre datos muy particulares que la herramienta solicita, como se indican en los ejemplos 1 y 2.

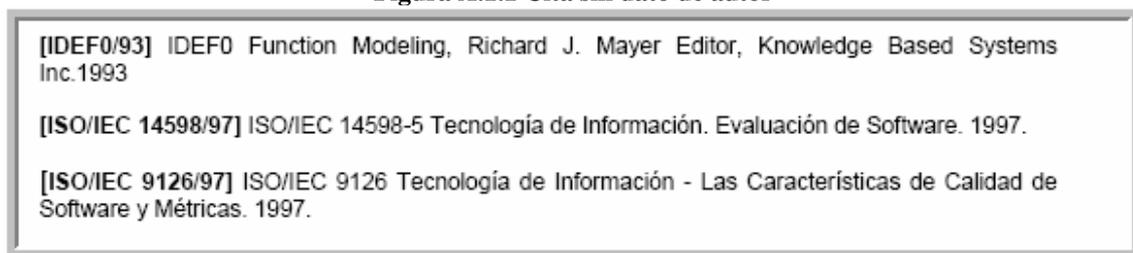
Ejemplo 1. En el caso de un artículo, autor (s), título, fuente, volumen, issue, inicio y fin de página, lugar de publicación y año. Algunos datos son opcionales como el mes, comentario, dirección, e-mail abstract y citas.

Ejemplo 2. En el caso de un libro, autor (s), título, fuente, lugar de publicación y año

Si por el contrario no se cuentan con estos datos, esta herramienta no los almacena. Durante el proceso de captura de datos hubo varios problemas que se detectaron al momento de hacer la captura de las citas en lo que respecta a los libros, y algunas de ellas se corrigieron haciendo una búsqueda por medio de un navegador de Internet, para así poder llevar a cabo el almacenamiento.

La Figura A.1.1 muestra cuando la cita indicada contiene un dato que no es un autor de la publicación.

**Figura A.1.1 Cita sin dato de autor**



**Fuente: Elaboración Propia con base de datos del CIC.2008**

Las Figuras A.1.2 y A.1.3 muestran cuando la cita no tiene año de publicación.

**Figura A.1.2 Cita sin año de publicación**

P.T. Brewer K. Lano. *Creating Specifications from Code: Reverse-Engineering Techniques*, Oxford University Computer Laboratory.

Fuente: Elaboración Propia con base de datos del CIC.2008

**Figura A.1.3 Otro ejemplo de cita sin año de publicación**

[1] Abraham Silberschatz. *Fundamentos de bases de Datos*. Mc. Graw Hill.

Fuente: Elaboración Propia con base de datos del CIC.2008

La Figura A.1.4 muestra cuando se trata de un artículo y no tiene número y volumen de publicación.

**Figura A.1.4 Cita sin número y volumen de publicación**

[ALBRETCH79] Albretch, A.J. "Measuring Application Development Productivity" Proc. IBM Application Development Symposium, Monterrey CA, Octubre 1979.

Fuente: Elaboración Propia con base de datos del CIC.2008

Y la Figura A.1.5 cuando no tiene editorial y año de publicación.

**Figura A.1.5 Cita sin editorial y año de publicación**

[Kogan, 1986] KOGAN, J.M. *Information for Motivation: A Key to Executive Information Systems that Translate Strategy Into Results for Management*.

Fuente: Elaboración Propia con base de datos del CIC.2008

Para tratar de corregir estas deficiencias se realizó lo siguiente, por ejemplo con los datos de un libro.

- 1) Por medio del buscador Google se buscaba el título de un libro
- 2) Se localizaban todos los datos del libro en búsqueda
- 3) Se formaba la correcta cita y se almacenaba en Excel
- 4) Al terminar de obtener los datos de todas las citas a libros de una tesis, estos se registraban en la herramienta informática Histcite Versión 8.5.26.

Lo anterior también se trató de desarrollar con los demás tipos de documentos indicados en la Tabla A.1.1, pero por el gran volumen de citas, se inclinó a trabajar solo con las citas a libros.

### ANEXO 3 CONCEPTOS DE INDICADORES DE CITACIÓN

La Tabla A.3.1 muestra los indicadores de análisis de citas que se deben de tomar en cuenta al aplicar un análisis cuantitativo Spinak (1988).

**Tabla A.3.1 Indicadores de Análisis de Citas**

No	Técnica	Medida
1	Activity index	Índice de actividad
2	Affinity index	Índice de afinidad
3	Attractivity index	Índice de atracción
4	Bibliographic coupling	Apareo bibliográfico
5	Citation analysis	Análisis de citaciones
6	Cocitation analysis	Análisis de cocitaciones
7	Consumption factor	Factor de consumo
8	Diversity index	Índice de diversidad
9	Echo factor	Factor de eco
10	Immediacy index	Índice de inmediatez
11	Impact factor	Factor de impacto
12	Impact index	Índice de impacto
13	Insularity index	Índice de aislamiento
14	Openness index	Índice de apertura
15	Popularity factor	Factor de popularidad
16	Self-citation rate	Índice de autocitación

**Fuente: Spinak.2008**

En las siguientes líneas se realiza una breve descripción sobre estos indicadores.

#### 1. Índice de Actividad

Permite caracterizar el esfuerzo relativo de un país en cierto campo de investigación, según se refleja en las publicaciones que produce (cociente que resulta el porcentaje de publicaciones del país en el campo de investigación, dividido por el porcentaje mundial de publicaciones del país en todas las disciplinas). (Sotolongo et al., 2001).

$$I_{act} = \frac{\% \text{ publicaciones del país en el campo de investigación}}{\% \text{ publicación mundial del país en todas las disciplinas}}$$

## 2. Índice de Afinidad:

Este indicador es usado para evaluar la tasa relativa de los intercambios científicos entre un país A y otro país B, en un periodo de tiempo, y en relación a toda la cooperación internacional entre estos dos mismos países sobre el mismo periodo. (Okubo, 1997).

La formula para calcular el Índice de Afinidad fue desarrollada por el Laboratorio de Evaluación de Prospectiva Internacional (LEPI) de Francia:

$$IA = \frac{COP(A-B)}{COP(A-WD)} \times 100$$

donde:

COP(A-B): representa el número de enlaces científicos (cooperación) entre A y B.

COP(A-WD): Es el número de ligas de cooperación entre A y el mundo.

## 3. Índice de Atracción:

El índice de atracción caracteriza la visibilidad relativa a partir del número de citas que atrae. (Zacca, 2010). Un valor de IA igual a 1 indica que las citas recibidas por un país se corresponden con la media regional. Un valor de IA mayor que 1 representa fortaleza en el país, ya que supera los dominios de referencia. Un valor menor que uno supone un incentivo para implementar mejoras en esta área.

$$I_{atr} = [(N_{citSP}(\text{pais}) / N_{citarea}(\text{pais}))] / [(N_{citSP}(\text{región}) / N_{citarea}(\text{región}))]$$

donde:

$N_{citSP}(\text{pais})$ : Número de citas de una especialidad en un país.

$N_{citarea}(\text{pais})$ : Número de citas de una área de investigación en un país.

$N_{citSP}(\text{región})$ : Número de citas de una especialidad en una región.

$N_{citarea}(\text{región})$ : Número de citas de una área de investigación en una región.

#### **4. Apareo bibliográfico**

De acuerdo a Spinak (1996), el apareo bibliográfico se produce cuando dos documentos hacen referencia por los menos a una misma publicación, lo que denota un signo de proximidad temática entre ambos documentos.

#### **5. Análisis de citación**

Sung y Jung (2008), las clasifica en seis áreas:

1. La naturaleza de las citas.
2. Autocitación.
3. Artículos, libros altamente citados.
4. Análisis sobre campos especiales. laboratorios, editoriales, lenguajes.
5. Usando aspectos: Se relaciona con aspectos sociales y culturales de como se llevan a cabo las citas
6. Cocitación o citas conjuntas: Si un artículo de un autor A y otro de un autor B, son citados conjuntamente por un tercer artículo.

#### **6. Análisis de cocitaciones**

De acuerdo a Sung & Jung (2008), este punto engloba el punto anterior.

#### **7. Factor de Consumo**

Yanovsky (1981), define el factor de consumo como el producto entre el factor de popularidad y el factor de citación.

$K_n = K_p * K_c$ , donde

$K_p$ : Factor de popularidad

$K_c$ : Factor de citación

#### **8. Índice de diversidad**

No se encontró información de este indicador.

**9. Factor de echo**

No se encontró información de este indicador.

**Índice de Inmediatez.**

El índice de inmediatez mide la rapidez con la cual es citado el "artículo promedio" de una revista. El Índice de Inmediatez se calcula dividiendo el número de citas a artículos publicados en un año dado, por el número de artículos publicados en ese mismo año.

***Índice de inmediatez***

Citas en el año 2002 a artículos publicados en el 2002 = 74

Número de artículos publicados en 2002 = 105

$$I_{\text{imm}} = \frac{\text{Citas a artículos del año}}{\text{Número de artículos publicados}} = \frac{74}{105}$$

Índice de inmediatez = 0,705

<http://www.uned.es/biblioteca/basesdedatos/jcr.pdf>

**11. Factor de Impacto**

El factor de impacto mide la frecuencia con la cual ha sido citado el "artículo promedio" de una revista en un año en particular. El factor de impacto de los JCR se calcula dividiendo el número de citas actuales de ítems publicados en los dos años anteriores, por el número total de artículos publicado en los dos años anteriores.

Ejemplo:

***Factor de impacto de una revista en el año 2003:***

Citas a artículos publicados en 2002 = 95    Total artículos publicados en 2002 = 106

Citas a artículos publicados en 2001 = 154    Total artículos publicados en 2001 = 92

Total (2002 + 2001) = 95 + 154 = 249    Total (2002 + 2001) = 106 + 92 = 198

$$\begin{array}{lcl} \text{Cálculo del factor de impacto} = & \frac{\text{Citas recibidas}}{\text{Artículos publicados}} & = \frac{249}{198} \end{array}$$

Factor de impacto = 1.258

<http://www.uned.es/biblioteca/basesdedatos/jcr.pdf>

## 12. Índice de impacto

Es el cociente entre el número de citas recibidas y el número de trabajos publicados. El índice de impacto relativo es el cociente entre el índice de impacto de un autor, revista o grupo y el índice de impacto máximo del campo al que pertenecen.

<http://informationr.net/ir/9-4/paper193.html>

## 13. Índice de aislamiento

Es el porcentaje de referencias que corresponden a publicaciones del mismo país de la publicación. (Escorcia, et al 2009)

$$I_{\text{aisl}} = \frac{\text{Citas}_{\text{país}}}{\text{TOTAL}_{\text{citas}}} \times 100 \%$$

## 14. Índice de apertura

Existen distintos índices o medidas que analizan diversos aspectos de las citas o referencias que hacen unos documentos a otros.

La tasa de auto citación muestra la proporción en que las citas de una revista se refieren a ella misma respecto al total de citas que hace. Esta tasa es una indicación del grado de clausura y apertura de la disciplina. La apertura de una revista puede verse bajo tres índices que varían entre cero y uno:

- La apertura total, que es el complemento del índice de auto citación.
- La apertura dentro de la misma disciplina

- La apertura fuera de la disciplina, que muestran cuan abierta es la revista a otras de la misma disciplina o de otros campos del conocimiento.

<http://www.bibliociencias.cu/gsd/collect/eventos/index/assoc/HASH013f/26f1bc40.dir/doc.pdf>

### **15. Factor de Popularidad**

Yanovsky (1981), define el Factor de Popularidad de la siguiente manera:

$KP = q1/q2$ , donde

$q1$  = Número de fuentes el cual son referenciadas dada una revista en un periodo

$q2$  = Número de fuentes el cual fueron citadas por una revista durante un periodo

### **16. Índice de autocitación**

Cuando un autor hace referencia a sus propias publicaciones, puede ser para poner en juego resultados anteriores o más prosaicamente para hacerse publicidad y muy a menudo por las dos razones a la vez. (Callon, Courtail y Penan, 1993: 25).

## **ANEXO 4. DESCRIPCIÓN DE LA METODOLOGÍA GENERAL PARA EL ANÁLISIS CUANTITATIVO**

La metodología a llevar a cabo en el análisis cuantitativo consta de los siguientes cinco pasos:

### **1. Definición de los objetivos y alcance de estudio**

Este punto, ya se describió en la introducción.

### **2. Selección de la fuente de datos y obtención de los datos**

El trabajo de esta tesis es aplicar el análisis cuantitativo en un área no analizada en México, la disciplina seleccionada fueron las ciencias de la computación, el generador de recursos de información es un Centro de Investigación, el cual pertenece al IPN e inicio sus actividades en 1996, ofreciendo maestría en la disciplina de las ciencias de la computación.

El recurso específico de información son los documentos de temas de tesis de maestría concluidas, a partir de 1996-2007, dando como resultado la colección de 198 tesis.

Los datos se recopilaron por información suministrada por la unidad de tecnologías educativas (control escolar), biblioteca del centro y por apoyo de profesores, los datos son los siguientes:

#### **a) Datos en las tesis:**

Nombre del autor de la tesis

Nombre de la tesis

Periodo del desarrollo de la tesis (año, mes y día)

Tipo de publicación

Lenguaje

País

Citas Referenciadas

Resumen

**b) Referenciada Bibliográficas en las tesis:**

Nombre (es) del Autor (es)  
Nombre del artículo, libro, conferencia  
Editorial  
Volumen  
Inicio y Fin de página  
Año, mes y día  
Tipo de publicación  
Lenguaje  
País

**3. Revisión a los valores de los datos del documento de tema de tesis**

El recurso o fuente de información base de estudio son las tesis de maestría, en lo que respecta a referencias bibliográficas y entre los que se encuentran citas a artículos, libros, congresos, notas técnicas, manuales y páginas de Internet. Sin embargo el análisis se realizó sobre citas a libros.

La decisión de solo analizar citas a libros se debió a tres aspectos importantes; el primero, existe una forma errónea de citar, segundo, las referencias bibliográficas se encontraban incompletas (año, editorial, autor, nombre de la obra) y el tercero los datos no eran los correctos; con el afán de realizar la corrección a estos tres aspectos, se hizo el ejercicio de hacer una revisión exhaustiva por Internet para encontrar los datos faltantes o corregir y así poder realizar el almacenamiento correcto de la referencia bibliográfica. Por la demanda de esfuerzo y tiempo requerido para realizar este ejercicio, se deja para un trabajo futuro y continuación de este trabajo de tesis.

**4. Tratamiento de los Datos**

Para el análisis de la Producción Científica, se usaron los siguientes datos:

1. Tema de tesis
2. Director
3. Codirector

4. Laboratorio
5. Investigadores Externos
6. Institución de procedencia de los alumnos

El detalle de este análisis, se puede ver en el capítulo 2.

Los datos de las citas a libros que se obtuvieron a partir de los documentos de tesis, se almacenaron primeramente en una base de datos en Excel, donde se hizo la depuración de los datos.

Para el análisis de citas se hace uso de una herramienta informática llamada HistCite versión 8.5.26. Este análisis se desarrolló y se muestra en el capítulo 2.

El proceso del análisis para generar los mapas tecnológicos, se muestra en el capítulo 3.

#### **5. Herramientas para visualizar los análisis de los datos**

Para el almacenamiento de los datos y las representaciones gráficas obtenidas a partir de ellos se llevó a cabo por medio de los paquetes informáticos Excel 2002, HistCite versión 8.5.26 y Redes 2005.

**ANEXO 5. VOCABULARIO EN LOS TÍTULOS DE LAS TESIS**

Rango	Frecuencia	Palabras/Descriptor	Rango	Frecuencia	Palabras/Descriptor	Rango	Frecuencia	Palabras/Descriptor
1	103		51	5	ASIGNACION	101	3	CONTEO
2	51	SISTEMA	52	5	BETA	102	3	DICCIONARIO
3	26	DATOS	53	5	DETECCION	103	3	EDIFICIOS
4	22		54	5		104	3	ESPACIALES
5	21	ANALISIS	55	5	ESTRUCTURAS	105	3	ESPACIO
6	19		56	5	MULTIAGENTE	106	3	GENERADOR
7	18		57	5	MUNDOS	107	3	IDENTIFICACION
8	17	MODELO	58	5	RECONOCIMIENTO	108	3	INDUSTRIALES
9	15	CONTROL	59	5	RECUPERACION	109	3	INTERFAZ
10	15	DESARROLLO	60	5	TECNICAS	110	3	LENGUAJE
11	14	SISTEMAS	61	5	VIRTUAL	111	3	MANEJO
12	13	DISEÑO	62	4	AGENTES	112	3	MEDIO
13	12	OBJETOS	63	4	ALGORITMO	113	3	MOMENTOS
14	12		64	4	APLICACIONES	114	3	MONITOREO
15	12	USANDO	65	4	AUTOMATIZACION	115	3	MOVIL
16	11	REDES	66	4	BASADA	116	3	NEURONALES
17	11	TIEMPO	67	4	CALIDAD	117	3	OPTIMIZACION
18	10	COMPONENTES	68	4	COMPORTAMIENTO	118	3	PLANTAS
19	10	SOFTWARE	69	4	CONTROLADOR	119	3	PRODUCCION
20	9	ESTUDIO	70	4	DISTRIBUIDAS	120	3	PROGRAMACION
21	9	EVALUACION	71	4	DISTRIBUIDO	121	3	ROBOT
22	9	IMAGENES	72	4	EDUCACION	122	3	SEMANTICAS
23	9	INFORMACION	73	4	GEOMETRICOS	123	3	SERIES
24	9		74	4	GESTION	124	3	SERVIDOR
25	8		75	4	INFRAESTRUCTURA	125	3	SIMILITUD
26	8	ESPAÑOL	76	4	INTERNET	126	3	TEXTOS
27	8	SIMULACION	77	4	JAVA	127	3	TRAFICO
28	8	WEB	78	4	MEDIANTE	128	3	
29	7	AMBIENTE	79	4	MEDICION	129	3	UBICUO
30	7	APLICACION	80	4	PALABRAS	130	3	VEHICULAR
31	7	ASOCIATIVAS	81	4	PLATAFORMA	131	3	VISUAL
32	7	BASES	82	4	RECURSOS	132	2	ACCESO
33	7	HERRAMIENTA	83	4	SERVICIO	133	2	AGREGADOS
34	7		84	4	SERVICIOS	134	2	AGRUPAMIENTO
35	7	PATRONES	85	4	TECNICA	135	2	ALTO
36	7	REAL	86	4	VRML	136	2	APLICADOS
37	7	VIRTUALES	87	3	ADMINISTRADOR	137	2	ASFALTENICOS
38	6	APRENDIZAJE	88	3	ADQUISICION	138	2	AUTOMATAS
39	6	AUTOMATICA	89	3	ALGORITMOS	139	2	BAJO
40	6	AUTOMATICO	90	3	APOYO	140	2	BASADOS
41	6	BASE	91	3	BIBLIOTECA	141	2	CALCULOS
42	6	IMPLEMENTACION	92	3	CALCULO	142	2	CAPACIDAD
43	6	MEMORIAS	93	3	CLASIFICADOR	143	2	CARACTERIZACION
44	6	METODOLOGIA	94	3	COLOCACIONES	144	2	CELULAR
45	6		95	3		145	2	COLABORATIVOS
46	6	RED	96	3	COMPARATIVO	146	2	COLOR
47	6	SOBRE	97	3	COMPLEJO	147	2	COMPILACION
48	6	UTILIZANDO	98	3	COMPUTACIONAL	148	2	COMPRESION
49	5	ALFA	99	3	COMUNICACION	149	2	COMPUTADORA
50	5	ARQUITECTURA	100	3	CONSTRUCCION	150	2	COMPUTADORAS

Rango	Frecuencia	Palabras/Descriptores	Rango	Frecuencia	Palabras/Descriptores	Rango	Frecuencia	Palabras/Descriptores
151	2	CONOCIMIENTO	201	2	PARALELA	251	1	APLICATIVOS
152	2	CONTEXTO	202	2	PLATAFORMAS	252	1	ARAÑAS
153	2	CONVERSION	203	2	PRESENCIA	253	1	ARBOLES
154	2	DECISIONES	204	2	PROCESO	254	1	ARBORESCENTES
155	2	DIFERENTES	205	2	GRAMAS	255	1	ARCHIVOS
156	2	DIGITAL	206	2	PUNTO	256	1	AREA
157	2	DIGITALES	207	2	REGISTROS	257	1	ARTESANALES
158	2	DISTRIBUIDOS	208	2	RELACIONES	258	1	ARTIFICIALES
159	2	DOCUMENTOS	209	2	RESOLUCION	259	1	ASEGURAMIENTO
160	2	DSP56827	210	2	SEGMENTACION	260	1	ASISTIDA
161	2	ELECTRONICA	211	2	SELECCION	261	1	ASISTIDO
162	2	EMPLEANDO	212	2	SEMIAUTOMATICA	262	1	ASOCIACION
163	2	ENTRADA	213	2	SIG	263	1	ASPECTOS
164	2	ESCRITORIO	214	2	SINTACTICAS	264	1	ASPERSION
165	2	ESPACIAL	215	2	SQL	265	1	ASSOCIATOR
166	2	ESTADISTICO	216	2	TELEFONIA	266	1	AUDITIVO
167	2	EXPERTO	217	2	TEXTO	267	1	AUDITIVOS
168	2	EXTENSION	218	2	TOMA	268	1	AUTENTICACION
169	2	EXTRACCION	219	2	TRABAJO	269	1	AUTENTICACION
170	2	FLUJO	220	2	TRIDIMENSIONALES	270	1	AUTOMATIZADA
171	2	FUENTES	221	2	VISUALIZACION	271	1	AUTOMOVILES
172	2	GENERACION	222	1	ABP	272	1	AUXILIARES
173	2	GENERALIZACION	223	1	ACEDER	273	1	BALANCEADOS
174	2	GENETICO	224	1	ACCIDENTES	274	1	BASADAS
175	2	GRAFICA	225	1	ACERO	275	1	BASICA
176	2	GRAFICO	226	1	ACTIVEX	276	1	BAYESIANAS
177	2	HARDWARE	227	1	ACTUALES	277	1	BILINGÜES
178	2	HERRAMIENTAS	228	1	ACUSTICO	278	1	BINARIO
179	2	IBM	229	1	ADMINISTRACION	279	1	BINARIOS
180	2	IMAGEN	230	1	ADMINISTRACIÓN	280	1	BIO
181	2	IMPRESAS	231	1	ADO	281	1	BIOINFORMATICA
182	2	INGENIERIA	232	1	ADORE	282	1	BODEGA
183	2	INTEGRACION	233	1	ADRAL	283	1	BOLETOS
184	2	INTEGRAL	234	1	AEROFONOS	284	1	BUSQUEDA
185	2	INTERPRETE	235	1	AFINES	285	1	C.C
186	2	INTRANET	236	1	AGENTE	286	1	CAMARA
187	2	JINI	237	1	AIRE	287	1	CAMARONICULTURA
188	2	KHEPERA	238	1	ALEATORIO	288	1	CAMBIOS
189	2	LERNMATRIX	239	1	ALUMBRADO	289	1	CAMINO
190	2	LINEALES	240	1	AMBIGÜEDAD	290	1	CAMPO
191	2	LOCAL	241	1	ANAFORA	291	1	CAPA
192	2	MAPAS	242	1	ANALITICO	292	1	CAPAS
193	2	MEXICANA	243	1	ANALIZADOR	293	1	CARACTERES
194	2	MINERIA	244	1	ANALIZANDO	294	1	CARACTERISTICAS
195	2	MODELADO	245	1	ANDERSON	295	1	CARRIL
196	2	MODELOS	246	1	ANIDADAS	296	1	CARTOGRAFICOS
197	2	MODULO	247	1	ANTE	297	1	CASE
198	2	MORFOLOGICA	248	1	ANTENA	298	1	CASO
199	2	MOVILES	249	1	APLICADO	299	1	CELULARES
200	2	ORIENTADOS	250	1	APLICANDO	300	1	CEREBRAL

Rango	Frecuencia	Palabras/Descriptor	Rango	Frecuencia	Palabras/Descriptor	Rango	Frecuencia	Palabras/Descriptor
301	1	CHIMENEAS	351	1	DCOM	401	1	ENCONTRAR
302	1	CICLO	352	1	DEFINICION	402	1	ENDS
303	1	CIENCIAS	353	1	DEFORMACIONES	403	1	ENERGIA
304	1	CLASES	354	1	DENTRO	404	1	ENRUTAMIENTO
305	1	CLASICAS	355	1	DESAMBIGUACION	405	1	ENSAYO
306	1	CLASIFICACION	356	1	DESCOMPOSICION	406	1	ENTIDADES
307	1	CLIENTE	357	1	DESCOMPRESION	407	1	EPC
308	1	CLINICA	358	1	DESCRIPCIONES	408	1	EQUIPO
309	1	CNC	359	1	DESCRIPTOR	409	1	ERRORES
310	1	CODIFICACION	360	1	DESPLIEGUE	410	1	ESCALA
311	1	CODIGO	361	1	DETERMINACION	411	1	ESCENAS
312	1	COLABORACION	362	1	D.I.S.A.D.A	412	1	ESCUELA
313	1	COLABORATIVO	363	1	DIAGNOSTICADOR	413	1	ESPACIOS
314	1	COLAS	364	1	DIAGRAMA	414	1	ESPANOL
315	1	COMBINACIONES	365	1	DIBUJO	415	1	ESPECTROSCOPIA
316	1	COMERCIO	366	1	DIFERENCIAS	416	1	ESQUEMA
317	1	COMPLEJOS	367	1	DIFUSO	417	1	ESTABILIZACION
318	1	COMPUTACIONALES	368	1	DIGITALIZACION	418	1	ESTACION
319	1	COMPUTARIZADO	369	1	DIMENSIONAMIENTO	419	1	ESTACIONARIAS
320	1	COMPUTO	370	1	DINAMICA	420	1	ESTACIONARIOS
321	1	COMUNICACIONES	371	1	DINAMICAMENTE	421	1	ESTACIONES
322	1	CONCEPCION	372	1	DINAMICO	422	1	ESTADISTICA
323	1	CONCEPTO	373	1	DIRECTA	423	1	ESTADO
324	1	CONCEPTOS	374	1	DIRECTORIOS	424	1	ESTANDAR
325	1	CONCEPTUALES	375	1	DISCRETOS	425	1	ESTIMACION
326	1	CONCEPTUALIZACION	376	1	DISEÑOS	426	1	ESTIMACIÓN
327	1	CONCURRENCIA	377	1	DISIMILARIDAD	427	1	ESTIMULOS
328	1	CONDICIONES	378	1	DISPOSITIVO	428	1	ESTOCASTICOS
329	1	CONDUCTO	379	1	DISPOSITIVOS	429	1	ESTRUCTURA
330	1	CONJUNTOS	380	1	DISTANCIA	430	1	ESTRUCTURADOS
331	1	CONSTRUCTOR	381	1	DISTRIBUCION	431	1	ESTÁNDAR
332	1	CONSULTA	382	1	DISTRIBUIDA	432	1	EVALUACIÓN
333	1	CONSULTAS	383	1	DISTRITO	433	1	EVENTOS
334	1	CONTAMINACION	384	1	DOMINIO	434	1	EVITANDO
335	1	CONTAMINANTES	385	1	DUPLICIDAD	435	1	EVOLUTIVA
336	1	CONTENIDOS	386	1	ECONOMICA	436	1	EXPLICATIVO
337	1	CONTINUA	387	1	EDICION	437	1	EXT
338	1	CONVERGENCIA	388	1	EDUCATIVO	438	1	EXTENDIDA
339	1	CORONARIOS	389	1	EDUCATIVOS	439	1	EXTENSIBLES
340	1	CORPUS	390	1	EFFECTIVA	440	1	EXTERNO
341	1	CORRIENTE	391	1	ELABORACION	441	1	FACILITAR
342	1	CORTO	392	1	ELECTRICA	442	1	FACTIBILIDAD
343	1	COSTOS	393	1	ELECTRICOS	443	1	FALLAS
344	1	CREACION	394	1	ELECTRONICAS	444	1	FARMACIAS
345	1	CREADOR	395	1	ELECTRONICO	445	1	FEDERAL
346	1	CREAR	396	1	ELEMENTOS	446	1	FENOMENOS
347	1	CRONOS	397	1	ELEVACION	447	1	FIGURAS
348	1	CUALIMETRICO	398	1	ELIMINAR	448	1	FINA
349	1	CUBICAS	399	1	EMISIONES	449	1	FIPA
350	1	CUIDADOS	400	1	ENCABEZADOS	450	1	FIRMA

Rango	Frecuencia	Palabras/Descriptorios	Rango	Frecuencia	Palabras/Descriptorios	Rango	Frecuencia	Palabras/Descriptorios
451	1	FLUJOS	501	1	INDUCCION	551	1	MAQUINAS
452	1	FONAVOZ	502	1	INESPERADOS	552	1	MARINAS
453	1	FONIATRIA	503	1	INFINITOS	553	1	MAS
454	1	FORMA	504	1	INFORMATICOS	554	1	MATEMATICA
455	1	FORMATEO	505	1	INMUNE	555	1	MATERIALES
456	1	FRACTALES	506	1	INSPIRADAS	556	1	MDCRNGM
457	1	FRACTALIDAD	507	1	INSTALACION	557	1	MEANS
458	1	FRACTURADAS	508	1	INSTALADO	558	1	MECA
459	1	FRAGMENTACION	509	1	INSTITUTO	559	1	MEDIDO
460	1	FRASEMAS	510	1	INSTRUMENTAL	560	1	MEJORAMIENTO
461	1	FRECUENCIA	511	1	INSTRUMENTO	561	1	MEJORAR
462	1	FRONTS	512	1	INTEGRIDAD	562	1	META
463	1	FUNCIONALIDADES	513	1	INTELIGIBILIDAD	563	1	METALICAS
464	1	FUNCIONES	514	1	INTERACCION	564	1	METODOLOGIAS
465	1	FUNCION	515	1	INTERACCIONES	565	1	METODOS
466	1	GENERAL	516	1	INTERFACES	566	1	METRICAS
467	1	GENETICOS	517	1	INTERFAMA	567	1	MEXICANO
468	1	GEOESPACIALES	518	1	INTERFASES	568	1	MEXICANOS
469	1	GEOGRAFICA	519	1	INTERIOR	569	1	MEXICO
470	1	GEOGRAFICO	520	1	INTERIORES	570	1	MICRONUCLEO
471	1	GEOMETRICAS	521	1	INTERMEDIA	571	1	MIDI
472	1	GEOMORFOLOGICOS	522	1	INTERPRETACION	572	1	MODELACION
473	1	GIS	523	1	INVARIANTES	573	1	MODIFICADO
474	1	GNU	524	1	JBDC	574	1	MODOS
475	1	GPRS	525	1	JDBC	575	1	MODULAR
476	1	GRAFICACION	526	1	JDMK	576	1	MODULARES
477	1	GRAFICACION	527	1	JOHNSON	577	1	MONITOR
478	1	GRAFOS	528	1	KERNEL	578	1	MONITORIZAR
479	1	GRAVIMETRICOS	529	1	KOHONEN	579	1	MONOVARIABLES
480	1	GRUPAL	530	1	LABERINTO	580	1	MONTADA
481	1	GRUPOS	531	1	LABORATORIOS	581	1	MORFOLOGICAS
482	1	GUIA	532	1	LEALTAD	582	1	MORFOLOGICO
483	1	HDT	533	1	LEXICOS	583	1	MOTIFS
484	1	HERENCIA	534	1	LINEA	584	1	MOTORES
485	1	HERRAMIENTAS	535	1	LINEAR	585	1	MOTORIZADA
486	1	HETEROGENEA	536	1	LINEAS	586	1	MPI
487	1	HETEROGENEAS	537	1	LINGÜÍSTICOS	587	1	MUESTREO
488	1	HIBRIDO	538	1	LINUX	588	1	MULTI
489	1	HISTORIA	539	1	LLAVE	589	1	MULTIAGENTES
490	1	HOMONIMIA	540	1	LOCALIZACION	590	1	MULTICRITERIO
491	1	HSI	541	1	M.S.ACCESS	591	1	MULTIDIMENSIONALES
492	1	HTML	542	1	MAGNETICA	592	1	MULTIESCALA
493	1	HUMANAS	543	1	MAGNETICO	593	1	MULTIMEDIA
494	1	HUMANOS	544	1	MAGNETOMETRICOS	594	1	MULTIPLES
495	1	I.M.P	545	1	MANEJADORES	595	1	MULTIVARIABLE
496	1	IDIOMA	546	1	MANIPULADOR	596	1	MULTIVARIADAS
497	1	IMPLEMENTADO	547	1	MANO	597	1	MódULO
498	1	INALAMBRICAS	548	1	MANTENIMIENTO	598	1	MóBIUS
499	1	INCREMENTAR	549	1	MAPPING	599	1	NATURALES
500	1	INDEPENDIENTES	550	1	MAQUINA	600	1	NECESARIAS

Rango	Frecuencia	Palabras/Descriptores									
601	1	NIVEL	651	1	PROBABILIDAD	701	1	ROBOTS	751	1	TELEOPERACION
602	1	NORMALIDAD	652	1	PROBLEMA	702	1	ROTATORIA	752	1	TEMPERATURA
603	1	NUCLEO	653	1	PROCESAMIENTO	703	1	RUIDO	753	1	TEMPORAL
604	1	NUOVA	654	1	PRODUCTOS	704	1	SABD	754	1	TEMPORALES
605	1	NUOVO	655	1	PROMEDIO	705	1	SALIDA	755	1	TERRENA
606	1	NUMERICA	656	1	PROPIETARIO	706	1	SAREO	756	1	TERTRIES
607	1	OBRA	657	1	PROPIESTO	707	1	SCMOD	757	1	TIPO
608	1	OBTENCION	658	1	PROTECCION	708	1	SECADORES	758	1	TOPOGRAFICO
609	1	ODBC	659	1	PROTCCOLOS	709	1	SEGUIMIENTO	759	1	TOPOGRAFICOS
610	1	ODICC	660	1	PROTOTIPOS	710	1	SEGURO	760	1	TOPOLOGICAMENTE
611	1	ODIO	661	1	PROYECTIVAS	711	1	SEMAFOROS	761	1	TOPOLOGICO
612	1	OLAP	662	1	PRUEBAS	712	1	SEMITONOS	762	1	TOSCA
613	1	OPERACION	663	1	PUBLICA	713	1	SENTIDOS	763	1	TPC
614	1	OPERACIONES	664	1	PUBLICACION	714	1	SERVER	764	1	TRANSACCIONES
615	1	OPERATIVO	665	1	PUNTES	715	1	SERVIDORES	765	1	TRANSFORMACIONES
616	1	OPERATIVOS	666	1	PUNTOS	716	1	SIMBOLICA	766	1	TRASLAPES
617	1	OPORTUNISTA	667	1	QNX	717	1	SIMBOLICOS	767	1	TRAYECTORIAS
618	1	ORB	668	1	QUAD	718	1	SIMPLES	768	1	TREES
619	1	ORGANIZACIONES	669	1	RADIACION	719	1	SIMPLIFICADAS	769	1	TRIANGULARES
620	1	ORIENTADA	670	1	RADIO	720	1	SIMULADO	770	1	TRIDIMENSIONAL
621	1	ORQUESTADOR	671	1	RASTER	721	1	SINCRONIZACION	771	1	TURBOCOMPRESORAS
622	1	OSCILOSCOPIO	672	1	RAT	722	1	SINTACTICA	772	1	UMBRAE
623	1	PACIENTE	673	1	RAZONAMIENTO	723	1	SINTETIZADOR	773	1	UNIDAD
624	1	PARALELO	674	1	RECINTOS	724	1	SOLUCION	774	1	UNIFICADO
625	1	PARAMETROS	675	1	RECUPERANDO	725	1	SONOROS	775	1	USO
626	1	PARQA	676	1	REDUCCION	726	1	SOPORTE	776	1	USUARIO
627	1	PARTICIONES	677	1	REFERENCIA	727	1	SP2	777	1	VAPORES
628	1	PARTICULAS	678	1	REFERENCIAL	728	1	SOLMX	778	1	VARIABLE
629	1	PATRON	679	1	REGIONALIZACION	729	1	SRSISTEM	779	1	VARIABLES
630	1	PERDIDAS	680	1	REGISTRO	730	1	STEINBUCH	780	1	VEHICULARES
631	1	PERFECTA	681	1	RELACIONAL	731	1	STENBUCH	781	1	VEHICULO
632	1	PERFILES	682	1	RELACIONALES	732	1	SUBMARINO	782	1	VELOCIDAD
633	1	PERSONAL	683	1	RELEVADOR	733	1	SUFICIENTES	783	1	VELOCIMETRIA
634	1	PERSONALIZADO	684	1	RELEVO	734	1	SUPERFICIES	784	1	VENTA
635	1	PETROLEO	685	1	REMOTA	735	1	SUPERSONICO	785	1	VIDA
636	1	PETROLERA	686	1	REMOTO	736	1	SUPERVISADA	786	1	VIDEOCONFERENCIA
637	1	PETROLERAS	687	1	RENDIMIENTO	737	1	SUPERVISADO	787	1	VIRUS
638	1	PETROLEROS	688	1	REPORTEADOR	738	1	SUS	788	1	VOICEXML
639	1	PLANTILLA	689	1	REPORTES	739	1	SWING	789	1	VOZ
640	1	PLASMAS	690	1	REPRESENTACION	740	1	TABU	790	1	WAVELETS
641	1	POLIFONIA	691	1	REPRESENTADOS	741	1	TAREAS	791	1	WBE
642	1	POSICIONAMIENTO	692	1	REPRESENTATIVO	742	1	TARJETA	792	1	XML
643	1	POZO	693	1	REPUBLICA	743	1	TARJETAS	793	1	YACIMIENTOS
644	1	PREALIMENTADO	694	1	RESOLVER	744	1	TCP	794	1	ZONAS
645	1	PREDICCION	695	1	RESONANCIA	745	1	TECNOLOGIA			
646	1	PREDICTIVAS	696	1	RESTAURACION	746	1	TEJIDO			
647	1	PRESION	697	1	RETINAS	747	1	TELE			
648	1	PREVENIR	698	1	REUSABLES	748	1	TELECONTROL			
649	1	PRIMARIOS	699	1	RFID	749	1	TELEFONICO			
650	1	PRIMITIVAS	700	1	RMI	750	1	TELEMONITOREO			

**ANEXO 6. ACM COMPUTING CLASSIFICATION SYSTEM. 1998 VERSION**

Como ejemplo se coloca de la clasificación el tópico original H, que corresponde a Sistemas de Información.

- \* H. Information Systems
  - o H.0 GENERAL
  - o H.1 MODELS AND PRINCIPLES
    - + H.1.0 General
    - + H.1.1 Systems and Information Theory (E.4)
      - # General systems theory
      - # Information theory
      - # Value of information
    - + H.1.2 User/Machine Systems
      - # Human factors
      - # Human information processing
      - # Software psychology NEW!
    - + H.1.m Miscellaneous
  - o H.2 DATABASE MANAGEMENT (E.5)
    - + H.2.0 General
      - # Security, integrity, and protection [\*\*]
    - + H.2.1 Logical Design
      - # Data models
      - # Normal forms
      - # Schema and subschema
    - + H.2.2 Physical Design
      - # Access methods
      - # Deadlock avoidance
      - # Recovery and restart
    - + H.2.3 Languages (D.3.2)
      - # Data description languages (DDL)
      - # Data manipulation languages (DML)
      - # Database (persistent) programming languages
      - # Query languages
      - # Report writers
    - + H.2.4 Systems
      - # Concurrency
      - # Distributed databases
      - # Multimedia databases NEW!
      - # Object-oriented databases NEW!
      - # Parallel databases NEW!
      - # Query processing
      - # Relational databases NEW!
      - # Rule-based databases NEW!
      - # Textual databases NEW!
      - # Transaction processing
    - + H.2.5 Heterogeneous Databases
      - # Data translation [\*\*]
      - # Program translation [\*\*]
    - + H.2.6 Database Machines

- + H.2.7 Database Administration
    - # Data dictionary/directory
    - # Data warehouse and repository NEW!
    - # Logging and recovery
    - # Security, integrity, and protection NEW!
  - + H.2.8 Database Applications
    - # Data mining NEW!
    - # Image databases NEW!
    - # Scientific databases NEW!
    - # Spatial databases and GIS NEW!
    - # Statistical databases NEW!
  - + H.2.m Miscellaneous
  - o H.3 INFORMATION STORAGE AND RETRIEVAL
    - + H.3.0 General
    - + H.3.1 Content Analysis and Indexing
      - # Abstracting methods
      - # Dictionaries
      - # Indexing methods
      - # Linguistic processing
      - # Thesauruses
    - + H.3.2 Information Storage
      - # File organization
      - # Record classification [\*\*]
    - + H.3.3 Information Search and Retrieval
      - # Clustering
      - # Information filtering NEW!
      - # Query formulation
      - # Relevance feedback NEW!
      - # Retrieval models
      - # Search process
      - # Selection process
    - + H.3.4 Systems and Software
      - # Current awareness systems (selective dissemination of information--SDI)
- [\*\*]
- # Distributed systems NEW!
  - # Information networks
  - # Performance evaluation (efficiency and effectiveness) NEW!
  - # Question-answering (fact retrieval) systems [\*\*]
  - # User profiles and alert services NEW!
- + H.3.5 Online Information Services
    - # Commercial services NEW!
    - # Data sharing
    - # Web-based services NEW!
  - + H.3.6 Library Automation
    - # Large text archives
  - + H.3.7 Digital Libraries (NEW!)
    - # Collection NEW!
    - # Dissemination NEW!
    - # Standards NEW!
    - # Systems issues NEW!
    - # User issues NEW!

- + H.3.m Miscellaneous
- o H.4 INFORMATION SYSTEMS APPLICATIONS
  - + H.4.0 General
  - + H.4.1 Office Automation (I.7)
    - # Desktop publishing NEW!
    - # Equipment [\*\*]
    - # Groupware NEW!
    - # Spreadsheets
    - # Time management (e.g., calendars, schedules)
    - # Word processing
    - # Workflow management NEW!
  - + H.4.2 Types of Systems
    - # Decision support (e.g., MIS)
    - # Logistics
  - + H.4.3 Communications Applications
    - # Bulletin boards
    - # Computer conferencing, teleconferencing, and videoconferencing
    - # Electronic mail
    - # Information browsers NEW!
    - # Videotex
  - + H.4.m Miscellaneous
- o H.5 INFORMATION INTERFACES AND PRESENTATION (e.g., HCI) (I.7)
  - + H.5.0 General
  - + H.5.1 Multimedia Information Systems
    - # Animations
    - # Artificial, augmented, and virtual realities
    - # Audio input/output
    - # Evaluation/methodology
    - # Hypertext navigation and maps [\*\*]
    - # Video (e.g., tape, disk, DVI)
  - + H.5.2 User Interfaces (D.2.2, H.1.2, I.3.6)
    - # Auditory (non-speech) feedback NEW!
    - # Benchmarking NEW!
    - # Ergonomics
    - # Evaluation/methodology
    - # Graphical user interfaces (GUI) NEW!
    - # Haptic I/O NEW!
    - # Input devices and strategies (e.g., mouse, touchscreen)
    - # Interaction styles (e.g., commands, menus, forms, direct manipulation)
    - # Natural language NEW!
    - # Prototyping NEW!
    - # Screen design (e.g., text, graphics, color)
    - # Standardization NEW!
    - # Style guides NEW!
    - # Theory and methods
    - # Training, help, and documentation
    - # User-centered design NEW!
    - # User interface management systems (UIMS)
    - # Voice I/O NEW!
    - # Windowing systems
  - + H.5.3 Group and Organization Interfaces

- # Asynchronous interaction
- # Collaborative computing NEW!
- # Computer-supported cooperative work NEW!
- # Evaluation/methodology
- # Organizational design
- # Synchronous interaction
- # Theory and models
- # Web-based interaction
- + H.5.4 Hypertext/Hypermedia (NEW!) (I.7, J.7)
  - # Architectures NEW!
  - # Navigation NEW!
  - # Theory NEW!
  - # User issues NEW!
- + H.5.5 Sound and Music Computing (NEW!) (J.5)
  - # Methodologies and techniques NEW!
  - # Modeling NEW!
  - # Signal analysis, synthesis, and processing NEW!
  - # Systems NEW!
- + H.5.m Miscellaneous (NEW!)
- o H.m MISCELLANEOUS