



Implantación práctica de un Modelo de Recuperación de Información Booleano

Indice

DESCRIPCIÓN	2
MODELO ENTIDAD-INTERRELACION DEL SISTEMA	2
MODELO RELACIONAL DEL SISTEMA.....	3
CONSULTAS EN SQL CONTRA EL SISTEMA DE RECUPERACIÓN	6



SISTEMAS AVANZADOS DE RECUPERACIÓN DE INFORMACIÓN

DESCRIPCIÓN

Se desea implementar un sistema de recuperación de Información basado en Tesoros de Descriptores en una base de datos Relacional. Los tipos de relaciones que queremos representar son: jerarquías, equivalencias y términos asociados.

El sistema debe ser capaz de almacenar todas las ocurrencias de un término del vocabulario en los documentos registrados en la base de datos.

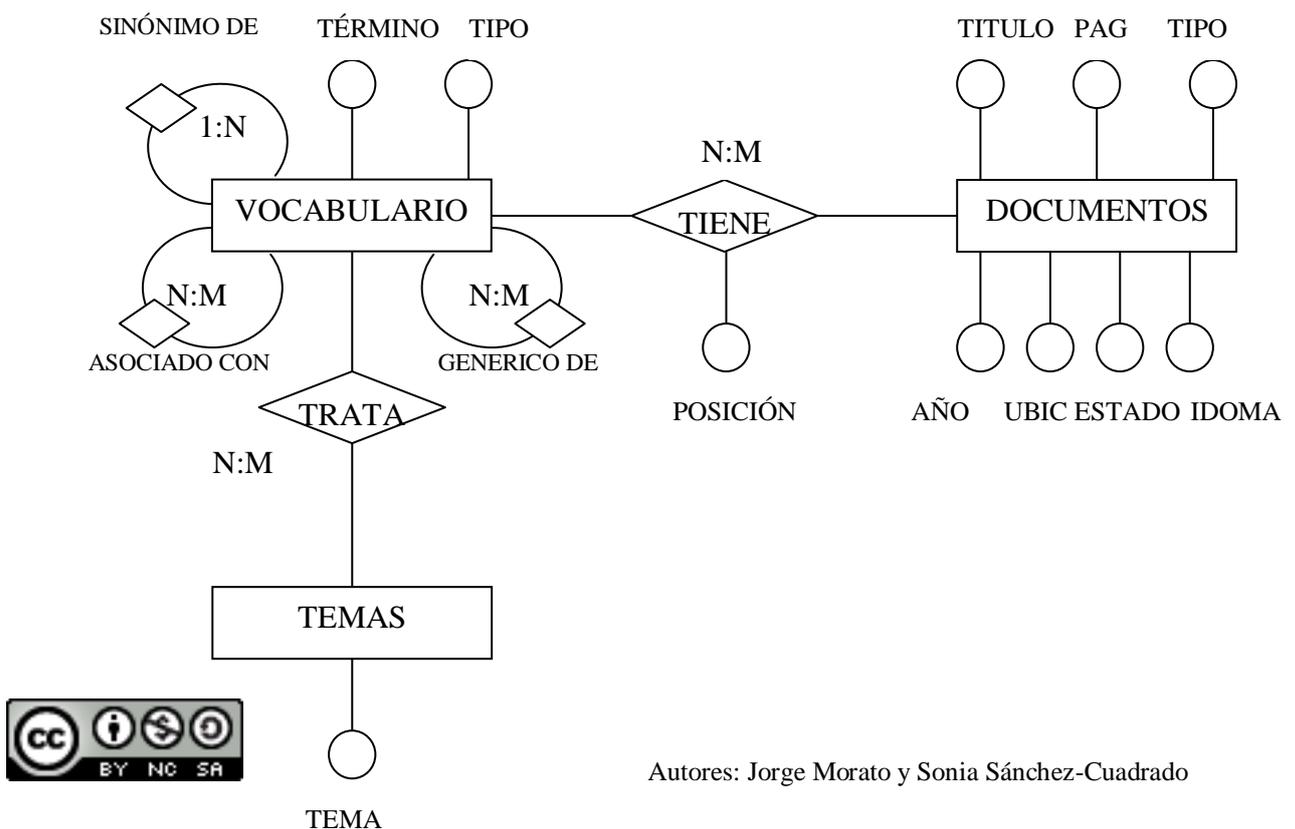
Un término genérico puede tener varios específicos y uno específico puede tener varios genéricos (polijerarquía).

En la base de datos, los términos relacionados no están influidos por el orden. Si el término A está relacionado con el término B, se da por supuesto que B está relacionado con A.

Un término del vocabulario puede tener varios temas asociados.

MODELO ENTIDAD-INTERRELACION DEL SISTEMA

Un posible modelo de entidad relación para el análisis del problema que nos ocupa puede ser el siguiente:



SISTEMAS AVANZADOS DE RECUPERACIÓN DE INFORMACIÓN

MODELO RELACIONAL DEL SISTEMA

Para pasar al modelo relacional deberemos de revisar las reglas de traslación:

- Por cada entidad se crea una tabla, cuyos campos sean los atributos del modelo E/R.
- Por cada relación 1:N, habrá que añadir un campo con la misma semántica de la clave principal de la entidad de ocurrencia 1 a la entidad que representa las N ocurrencias. Ese nuevo campo se denomina clave ajena y servirá para relacionar las dos tablas correspondientes.
- Por cada relación N:M, se crea una nueva tabla cuyos campos son las claves principales de las entidades que relaciona.

Las relaciones reflexivas siguen en realidad las reglas anteriores, pero por ser en realidad las dos entidades que rodean a la interrelación una única entidad, el número de tablas es menor:

- En una relación reflexiva 1:N, como es el caso de los sinónimos, se añade la clave ajena a la misma tabla que tiene la clave principal.
- En una relación N:M, los dos lados de la relación son la misma entidad, por lo que como resultado sólo habrá una única tabla, en vez de dos, por la parte de entidades. Por la relación, habrá que hacer otra tabla con la clave principal y la clave ajena de la entidad relacionada.

TABLA TDVOCABULARIO

COD_T	TIPO	DESCRIPTOR	TERMINO
1	1	1	RADAR
2	1	2	SONAR
3	1	3	EQUIPOS ELECTRONICOS
4	1	4	RAD86-34
5	1	1	RAD-AR
6	1	5	ELECTRONICA

TABLA TDGEN_ESP

COD_T	COD_GEN
1	3
2	3
3	6
4	1

TABLA TDRELACIONES





SISTEMAS AVANZADOS DE RECUPERACIÓN DE INFORMACIÓN

COD_T	COD_R
1	2

TABLA TDDOCUMENTOS

COD_DOC	TIPO	TITULO	AUTOR	UBICACION	PAG	DOCUMENTO	ESTADO	IDIOMA	TIPO_DOC	HYPERLINK	AÑO
1	1	MANUAL DE	PACO	LEG	250		PRE		MANUAL	http://www.u	2001
2	1	MANUAL DE	LUIS	LEG	720				MANUAL	http://www.y	1996
3	1	SISTEMA	PACO	GET	800		PRE		DOC_TRAB	http://www.el	1998
4	1	MANUAL DE	JAIME	GET	200				MANUAL	http://www.te	2000
5	1	SISTEMAS	PEDRO	COL	325					http://www.el	1994
7	1	SISTEMA	JUAN	LEG	125					http://algarro	1983
8	1	SISTEMA	JAVIER	LEG	234		PRE		DOC_TRAB	http://algarro	1998

TABLA TDREFERENCIAS

COD_T	COD_DOC	POSICION
1	1	5
1	1	38
1	1	69
1	1	85
2	2	8
2	2	56
2	2	79
2	2	98
3	1	9
3	1	76
3	1	109
3	2	79
3	2	100
3	2	134
3	5	4
3	5	42
3	5	56
3	5	98
4	3	13
4	3	28
4	3	65
4	3	98
4	3	130



SISTEMAS AVANZADOS DE RECUPERACIÓN DE INFORMACIÓN

4	4	23
4	4	40
4	4	85
4	4	98
4	4	140
5	5	2
5	5	36
5	5	46
5	5	58
5	5	100
6	3	15
6	3	32
6	3	79
6	3	95
6	3	128
6	4	20
6	4	36
6	4	76
6	4	89
6	4	137
1	8	125
2	8	64

TABLA TDATT

COD_A	AREA
1	ELECTRONIC
2	TIPOS DE

TABLA TDATT_AFINES

COD_A	COD_AFIN
1	2

TABLA TDATTCONCEPTOS

COD_T	COD_A
1	1
2	1
3	1
5	1
6	1
1	2
4	2
5	2



SISTEMAS AVANZADOS DE RECUPERACIÓN DE INFORMACIÓN

CONSULTAS EN SQL CONTRA EL SISTEMA DE RECUPERACIÓN

1. Selecciona los documentos que hayan sido escritos por Paco.
2. Seleccionar de la tabla documentos aquellos registros en que el nombre del autor comience por la letra P y cuyo ejemplar se encuentre prestado
3. Seleccionar de la tabla documentos aquellos registros cuyo nombre de autor contenga la letra "e".
4. Seleccionar los campos ubicación y páginas de la tabla documentos siempre y cuando los registros que se muestren sean del tipo manual y su autor sea Paco y estén en Getafe
5. Seleccionar de la tabla documentos los registros cuyo nombre de autor finalice con la letra "e" . Sólo deberás seleccionar los campos título, tipo y autor. Además, los registros seleccionados deberán contener entre las palabras del título la palabra "manual" o la palabra "radar"
6. Seleccionar de la tabla documentos los registros que contengan la palabra "radar" en el título o aquellos registros que cumplan al mismo tiempo las siguientes condiciones: que el nombre del autor finalice con la letra "e" y que contengan la palabra "manual" en el título.
7. Seleccionar de la tabla documentos aquellos registro cuyo título contenga la palabra manual y cuyo tipo sea o bien manual o bien documento de trabajo.
8. Selecciona de la tabla documentos los registros en los que el campo estado esta en blanco.
9. Documentos que tengan 500 páginas o más.
10. Selecciona los documentos que no contengan la palabra "manual" en el título. Muestra los campos título y páginas. Ordena los registros de forma decreciente por el campo página.
11. Selecciona el campo título de la tabla documentos de aquellos registros cuyo año se encuentre entre el 1996 y el 1999, ambos inclusive. Escribe esta consulta de dos maneras diferentes.
12. Muestra los documentos que tengan menos de 500 hojas y que no hayan sido escritos por Paco. Muestra los campos título y autor, ordenándolo de manera ascendente por el campo autor
13. Muestra el campo titulo de la tabla documentos seleccionando los registros en los el campo estado no este vacío.
14. Muestra los registros de la tabla documentos que al mismo tiempo ni estén escritos por Paco ni estén en Getafe. Muestra los campos página y estado, ordena los registros de forma decreciente por el campo página





SISTEMAS AVANZADOS DE RECUPERACIÓN DE INFORMACIÓN

16. Muestra los registros que tienen o bien entre 200 y 600 páginas o bien el año es mayor que el 1998. Seleccionando además aquellos que contengan en el título las letras "de".

18. Muestra los autores diferentes que hay en el sistema sin utilizar el group

REALIZAR CALCULOS

Elementos: SUM, AVG, *, +, -, /, MIN, MAX, COUNT

19. Suma 7 al campo año (recuerda que año debe de ser numérico). Muestra el campo año y el campo nuevo en el que aparece el año más siete. Llama al campo que has creado SUMA.

20. Suma siete tercios al número de hojas.

21. Calcula la media de páginas de todos los documentos de la base de datos.

22. Muestra el año menor de todos los documentos de la base de datos.

23. Muestra cual es el último libro de cada autor

24. Suma el numero de hojas de todos los documentos

25. Suma el numero de hojas de los documentos que tienen como tipo manual

26. Cuenta el número de documentos que ha escrito Paco

27. Cuenta el número de documentos que ha escrito cada autor

28. Cuenta el número de manuales que ha escrito cada autor.

29. Cuenta los títulos de los documentos distintos que hay en la BD (dos documentos podrían considerarse diferentes aún teniendo el mismo título, ya que pueden referirse a ediciones distintas).

SELECTS CONTRA EL TESAURO!

30. Descriptores por los que se ha indizado el documento Manual de radares

31. Genéricos de radar

32. Específicos de radar

33. Términos relacionados con radar

34. Genéricos y específicos de radar

35. Sinónimos de radar

36. Documentos que traten del descriptor radar





SISTEMAS AVANZADOS DE RECUPERACIÓN DE INFORMACIÓN

37. Documentos que traten de los descriptores radar O sonar
38. Misma consulta pero con radar Y sonar
39. Descriptores comunes entre manual de radares y manual de sonares
40. Documentos que traten del descriptor RADAR o sus sinónimos
41. Documentos que traten del descriptor RADAR o de sus genéricos
42. Documentos que traten del descriptor RADAR o sus específicos
43. Documentos que traten del descriptor RADAR o sus asociados
44. Documentos que contengan términos que traten del tema 'ELECTRÓNICA'
45. Documentos que contengan términos que traten de los temas relacionados con el tema 'ELECTRONICA'
46. Numero de ocurrencias del descriptor RADAR en el documento "MANUAL DE RADARES"
47. Documentos que contengan el descriptor "EQUIPOS ELECTRONICOS" ordenados por nº de ocurrencias
48. Lista los términos indizados del documento "MANULA DE RADARES" por orden de frecuencia de aparición
49. documentos que contengan RADAR Y SONAR separados a menos de 5 posiciones
50. Documentos que contengan la frase "sonar radar"

