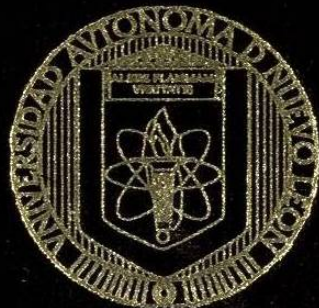


UNIVERSIDAD AUTONOMA DE NUEVO LEON  
FACULTAD DE CONTADURIA PUBLICA  
Y ADMINISTRACION



LOS DOCUMENTOS ELECTRONICOS:  
MODELIZACION Y CONSULTA DE DOCUMENTOS  
ESTRUCTURADOS BAJO UN AMBIENTE  
ORIENTADO A OBJETOS.

POR  
ABRAHAM ALVAREZ ESCOBEDO

COMO REQUISITO PARCIAL PARA OBTENER  
EL GRADO DE MAESTRIA EN INFORMATICA  
ADMINISTRATIVA

DICIEMBRE, 2001

TM

Z7164

.C8

FCPYA

2001

.A48



1020145627

UNIVERSIDAD AUTONOMA DE NUEVO LEON  
FACULTAD DE CONTADURIA PUBLICA  
Y ADMINISTRACION



LOS DOCUMENTOS ELECTRONICOS:  
MODELIZACION Y CONSULTA DE DOCUMENTOS  
ESTRUCTURADOS BAJO UN AMBIENTE  
ORIENTADO A OBJETOS.

POR  
ABRAHAM ALVAREZ ESCOBEDO

COMO REQUISITO PARCIAL PARA OBTENER  
EL GRADO DE MAESTRIA EN INFORMATICA  
ADMINISTRATIVA

DICIEMBRE, 2001

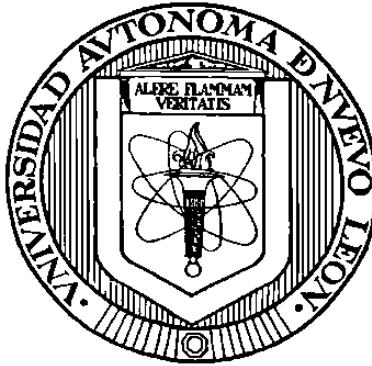
975402.

TM  
27164  
.C8  
FOPYA  
2001  
.A48



FONDO  
TESIS

**UNIVERSIDAD AUTONOMA DE NUEVO LEON  
FACULTAD DE CONTADURIA PUBLICA Y  
ADMINISTRACION**



**LOS DOCUMENTOS ELECTRONICOS:  
MODELIZACION Y CONSULTA DE DOCUMENTOS  
ESTRUCTURADOS BAJO UN AMBIENTE ORIENTADO A  
OBJETOS.**

**Por**

**ABRAHAM ALVAREZ ESCOBEDO**

**Como requisito parcial para obtener el Grado de  
MAESTRIA EN INFORMATICA ADMINISTRATIVA**

**Diciembre, 2 0 0 1**

**LOS DOCUMENTOS ELECTRONICOS:  
MODELIZACION Y CONSULTA DE DOCUMENTOS  
ESTRUCTURADOS BAJO UN AMBIENTE ORIENTADO A  
OBJETOS.**

**Aprobación de la Tesis :**

---

**M.A. Jesús Fabian López Pérez**  
Asesor de la Tesis

---

**C.P. Francisco J. Jardines Garza**  
Coordinador de Postgrado

# Agradecimientos

" I was taught that the way of progress  
is neither swift nor easy"

*Marie Curie*

DOY GRACIAS ...

*A Dios, por la vida que me ha dado ...*

*A mis padres, por todo su gran sacrificio*

*A mis Hermanos por su apoyo y sustento*

*A mis Maestros por su tiempo y dedicación*

*y*

*A Verónica por todo su amor ...*

---

Al Gobierno de México, a través del Consejo Nacional de Ciencia y  
Tecnología, por el soporte financiero.



# Tabla de Contenido

Capítulo	Página
<b>1 Estado del Arte</b>	<b>1</b>
1.1 Introducción	1
1.2 Problemática	3
1.3 Alcance	4
1.4 Objetivos de trabajo	4
1.5 Organización de la tesis	5
<b>2 Sistemas de Búsqueda de Información (SBI)</b>	<b>7</b>
2.1 Introducción	7
2.2 Los diferentes tipos de búsqueda	8
2.3 Los SBI y los SMBD	9
2.4 Funcionalidades de un SBI	10
2.4.1 Indexación	11
2.4.2 Interrogación	12
2.5 Modelos de búsqueda de información	13
2.5.1 El modelo boleano	13
2.5.2 El modelo vectorial	14
2.5.3 El modelo probabilista	15
2.5.4 El modelo lingüístico	15
2.6 Modos de Búsqueda	15
2.6.1 Los SMBD	16
2.6.2 Los SBI	16
2.6.3 Los Sistemas Hypertextos	17
2.6.4 Ejemplos de sistemas hypertextos	18
2.6.5 Los sistemas hypermedia	19
2.6.6 La diferencia entre un SBI y un SGBD	19
2.7 Enfoques de los modelos conceptuales	20
2.8 Clasificación de los sistemas de búsqueda de información	21
2.8.1 Estructuras de archivos	21
2.8.2 Enfoque archivo plano	21
2.8.2.1 Algoritmo básico	22
2.8.3 Enfoque archivo en sentido inverso	22
2.8.4 Enfoque de la firma	24
2.9 Conclusión	24

Capítulo	Página
<b>3 Normas de representación de los documentos estructurados</b>	<b>25</b>
3.1 Introducción	25
3.1.1 Principio	26
3.2 El modelo SGML	28
3.2.1 Los documentos SGML : estructuras específicas	28
3.2.2 La Definición de Tipo de Documento	30
3.2.3 Los elementos SGML	30
3.2.4 Los atributos SGML	32
3.2.5 Las entidades	34
3.2.6 Ejemplo simple de un DTD	35
3.2.7 La estructura Lógica	38
3.2.8 La estructura física	38
3.3 Conclusión	38
<b>4 Sistemas para la consulta de documentos estructurados</b>	<b>40</b>
4.1 Maestro : Un sistema de consulta	40
4.1.1 El lenguaje de consultas MAESTRO	41
4.1.2 Consultas sobre el contenido y sobre la estructura	43
4.1.3 Consultas sobre los atributos y caminos de acceso	44
4.2 MULTOS : Un sistema de consulta	45
4.2.1 Sintaxis para realizar consultas con MULTOS	47
4.3 LOREL: Un sistema de consulta	48
4.3.1 LOREL : Un sistema de consulta	48
4.4 Sistema de consulta Textriever	51
4.4.1 Modelo formal	52
4.5 POQL : Un sistema de consulta	53
4.5.1 Los operadores de POQL	53
4.5.2 Consultas sobre el contenido textual	54
4.5.3 Consultas sobre las uniones de tipos	55
4.6 Conclusión	56

Capítulo	Página
<b>5 Modelización</b>	<b>58</b>
5.1 Introducción	58
5.2 El modelo	59
5.2.1 El acceso sobre la estructura	59
5.2.2 El acceso sobre los atributos	63
5.2.3 El acceso sobre el contenido semantico	64
5.3 Covertura y dependencias	65
5.4 La función de los atributos	67
5.5 Proceso de propagación de los atributos portadores	70
5.6 Indexación estructural	71
5.7 Implementación y contribuciones al modelo	73
5.7.2 La organización de los atributos	73
5.7.3 Acceso por consultas	75
5.7.4 Acceso a través de la navegación	76
5.8 Otras contribuciones	76
5.8.1 Características	77
5.8.2 Semantica	77
5.8.3 Dominios	78
5.8.4 Tipos y Clases	78
5.8.5 Valores y Objetos	79
5.9 Conclusión	81
<b>6 Presentación del Prototipo</b>	<b>83</b>
6.1 Introducción	83
6.1.1 Presentación del prototipo	83
6.2 Características del sistema O2	84
6.2.1 Esquemas, tipos y clases	84
6.2.2 Objetos, Valores y objetos nombrados	85
6.3 La arquitectura	86
6.4 El manipulador de documentos	89
6.5 La base de documentos	90

<b>Capítulo</b>	<b>Página</b>
<b>7 Conclusión General</b>	<b>92</b>
7.1 Perspectivas	93
<b>Bibliografía</b>	<b>94</b>
<b>Lista de Cuadros</b>	<b>98</b>
<b>Lista de Figuras</b>	<b>99</b>