



**ESCUELA SUPERIOR POLITECNICA DEL
CHIMBORAZO**

FACULTAD DE INFORMATICA Y ELECTRÓNICA

ESCUELA DE INGENIERIA EN SISTEMAS

**“DESARROLLO DE UN PORTAL SEMÁNTICO BAJO EL PARADIGMA
XML TOPIC MAPS (XTM): CASO PRÁCTICO LICEO STEPHEN HAWKING”**

TESIS DE GRADO

Previa a la obtención del título de

INGENIERO EN SISTEMAS INFORMÁTICOS

Presentado por:

Mónica Pamela Fuertes Ortega

RIOBAMBA – ECUADOR

2009

AGRADECIMIENTOS

Deseo me permitan agradecer primeramente a Dios, por la suerte que tuve para superarme.

Expresar el sentimiento de gratitud y reconocimiento a la institución que me abrió las puertas Escuela Superior Politécnica de Chimborazo ESPOCH, a la Facultad de Informática y Electrónica, Escuela de Ingeniería en Sistemas, principalmente a mis profesores Ing. Raúl Rosero DIRECTOR DE TESIS, Ing. Jorge Huilca, ASESOR PRINCIPAL, maestros que con dedicación y sacrificio me guiaron por la ruta del saber.

Porque a más de cumplir con la delicada responsabilidad de educar inculcando los principios del deber, del honor la justicia y la solida moral, han trabajado incansablemente por mejorar las condiciones físicas y pedagógicas de este sagrado patrimonio, que constituye un verdadero templo de la educación, la ciencia y la cultura de nuestros pueblos.

Agradecer del mismo modo al Liceo de Talentos Stephen Hawking, lugar donde me permitieron realizar esta investigación y a todas las personas y amigos que me ayudaron e hicieron posible la culminación de este trabajo.

INFINITAMENTE GRACIAS A TODOS

Mónica Pamela Fuertes Ortega

DEDICATORIA

Cuando, uno realiza algo pensando en el bienestar de su familia, y motivado por el cariño, ternura y apoyo de quienes ama el esfuerzo es mayor y todo sacrificio vale la pena, para lograr alcanzar sus ideales.

A todas aquellas personas que aportaron en la consecución de esta aspiración profesional.

Mónica Pamela Fuertes Ortega

FIRMAS DE RESPONSABILIDAD

NOMBRE	FIRMA	FECHA
Dr. Ms.C. Romeo Rodríguez DECANO DE LA FACULTAD DE INFORMÁTICA Y ELECTRÓNICA
Ing. Iván Menes DIRECTOR DE LA ESCUELA DE INGENIERÍA EN SISTEMAS
Ing. Raúl Rosero DIRECTOR DE TESIS
Ing. Jorge Huilca MIEMBRO DEL TRIBUNAL
Lcdo. Carlos Rodríguez DIRECTOR DEL CENTRO DE DOCUMENTACIÓN
NOTA DE LA TESIS	

“Yo, Mónica Pamela Fuertes Ortega, soy responsable de las ideas, doctrinas y resultados expuestos en esta tesis; y, el patrimonio intelectual de la Tesis de Grado pertenece a la ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO.”

Mónica Pamela Fuertes Ortega

ÍNDICE DE ABREVIATURAS

ABREVIATURAS	SIGNIFICADO
ANSI	Abstract Syntax Notation1
ASCII	<i>American Standard Code for Information Interchange</i>
dHTML	Dynamic HTML
DOM	Document Object Model (Modelo de Objeto Documento)
DTD	Document Type Definition
ESPOCH	Escuela Superior Politécnica de Chimborazo
HTML	HyperText Markup Language(Lenguaje de Marcas de Hipertexto)
IIS	Servicio de información internet (Internet Information Services)
NLP	Natural Language Processing (<i>Procesamiento de Lenguajes Naturales</i>)
SGML	<i>Standard Generalized Markup Language</i>
SQL	Lenguaje estructurado de consulta (Structured Query Language)
SRS	Especificación de requerimientos de software
SVG	<i>Scalable Vector Graphics</i>
TAO	Topics, Asociaciones, Ocurrencias
TEI	Text Encoding Initiative

TM	Topic Maps
VML	<i>Vector Markup Language</i>
W3C	World Wide Web Consortium
XHTML	eXtensible Hypertext Markup Language (lenguaje extensible de marcado de hipertexto)
XML	Lenguaje de marca extensible (Extensible Markup Language)
XSL	Extensible Stylesheet Language
XSLT	Lenguaje de estilos extendido transformado (Extensible Style Language Transformaciones)
XTM	XML Topic Maps

INDICE

CAPÍTULO I

MARCO DE REFERENCIA.....	22
1.1. Título de la investigación.....	22
1.2. Problema de la Investigación.....	22
1.2.1. Análisis	23
1.2.2 Limitación.....	23
1.3. Objetivos.....	23
1.3.1. Objetivo General.....	23
1.3.2. Objetivos Específicos	23
1.4. Justificación de la Investigación.....	24
1.4.1 Justificación Teórica.....	24
1.4.2 Justificación Práctica	24
1.5. Hipótesis	25

CAPÍTULO II

MARCO TEORICO	26
2.1. Representación del conocimiento.....	26
2.1.1. Ontologías.....	26
2.1.2. Tesauros	28

2.1.3. Topic Maps	30
2.2. Representación de la información en la web	36
2.2.1. SGML (Standard Generalized Markup Language).....	36
2.2.2. Tecnología HTML (Hypertext Markup Language)	41
2.2.3. Tecnología XML (eXtensible Markup Language)	44
2.2.4. Tecnología XHTML (eXtensible Hypertext Markup Language).....	49
2.3. Representación del conocimiento bajo Topic Maps TM.....	53
2.3.1. Evolución de TM	54
2.3.2. Definición de TM.....	55
2.3.3. Modelos de TM.....	61
2.3.3.1. TMQL (Topic Maps Query Language)	61
2.3.3.2. TMCL (Topic Maps Constraint Language).....	62
2.3.3.3. XTM (XML Topic Maps).....	62
2.3.3.4. HyTM.....	63
2.3.4. Funcionamiento	64
2.3.5. Ventajas y Desventajas	66
 CAPÍTULO III	
REPRESENTACIÓN DEL CONOCIMIENTO BAJO XML TOPIC MAPS XTM	68
3.1. Evolución de XTM	68
3.2. Definición de XTM.....	70

Objetivos.....	71
Documento XTM.....	71
Limitación de denominación de tópico.....	71
Tópico	72
Ocurrencia.....	73
Asociación	74
3.3. Funcionamiento de XTM.....	75
Introducción a la sintaxis XTM	75
3.4. Ventajas y Desventajas de XTM	76
3.5. Normativas actuales	77
XTM 1.0 Definición del Tipo de Documento (Normativa).....	77
<i>Figura III.1.DTD del XTM</i>	82
3.6. Software	83
CAPÍTULO IV	
DESARROLLO E IMPLEMENTACIÓN DEL PORTAL SEMÁNTICO	84
4.1. FASE DE ANÁLISIS	85
4.1.1. EXPLORACION DEL ÁMBITO Y DEL PROBLEMA	85
4.1.2. FACTIBILIDAD DEL PROYECTO.....	98
4.1.3. ESPECIFICACIÓN DE REQUERIMIENTOS DE SOFTWARE (SRS)	101
4.2 FASE DE DISEÑO.....	107

4.2.1. PRESUPUESTO DEL COSTO POR EL ESFUERZO DE DESARROLLO	
.....	107
4.2.2. PLANIFICACIÓN Y SEGUIMIENTO DEL PROYECTO.....	116
4.2.3. GESTIÓN DE RIESGOS.	120
4.2.4. REFINAMIENTO DE LAS HISTORIAS DE USUARIOS	129
4.2.5. DIAGRAMA DE CLASES (MODELO CONCEPTUAL)	135
4.2.6. DIAGRAMAS DE CALLE	137
4.3. FASE DE CODIFICACIÓN.....	139
4.3.1. CAPA DE ACCESO A DATOS	139
4.3.2. CAPA DE APLICACIÓN E INTERFAZ	145
4.3.4. MODELADO FISICO	153
4.3.5. DIAGRAMA DE COMPONENTES.....	154
4.4. FASE DE PRUEBAS	156
4.4.1. PRESENTACIÓN DE LOS TEST FINALIZADOS.....	156
ANÁLISIS DE RESULTADOS DE LA HIPÓTESIS	161
ESTADISTICA INFERENCIAL	161
SELECCIÓN Y DETERMINACIÓN DE LA MUESTRA	161
OBTENCIÓN DE LOS DATOS	161
CLASIFICACIÓN Y ORGANIZACIÓN DE LOS DATOS	163
ANÁLISIS DE LOS DATOS	164

REPRESENTACIÓN GRÁFICA DE LOS RESULTADOS..... 168

VALIDACIÓN DE LA HIPÓTESIS..... 169

CONCLUSION DEL ANÁLISIS DE LA HIPÓTESIS 170

CONCLUSIONES

RESUMEN

SUMMARY

BIBLIOGRAFÍA

ANEXOS

Encuesta a usuarios

ÍNDICE DE TABLAS

Tabla II.1. Características de las diferentes formas de Representación del Conocimiento	36
Tabla II.2. Características de los diferentes tipos de representación de la información en la web.....	53
Tabla IV.1. Historia de usuario1 del sistema actual	92
Tabla IV.2. Historia de usuario2 del sistema actual	93
Tabla IV.3. Historia de usuario3 del sistema actual	93
Tabla IV.4. Historia de usuario4 del sistema actual	94
Tabla IV.5. Historia de usuario5 del sistema actual	94
Tabla IV.6. Historia de usuario1 del sistema propuesto	95
Tabla IV.7. Historia de usuario2 del sistema propuesto	95
Tabla IV.8. Historia de usuario3 del sistema propuesto	96
Tabla IV.9. Parámetros a satisfacer por el portal.....	99
Tabla IV.10. Software Reutilizable	100
Tabla IV.11. Costo de Sistema Propuesto	100
Tabla IV.12. Total Puntos de Función.....	111
Tabla IV.13. Factor costo Puntos de Función.....	111
Tabla IV.14. Factores de ponderación de complejidad	112

Tabla IV.15. Modelo Constructivo de Costo según Cocomo.....	113
Tabla IV.16. Costo de trabajador por mes	113
Tabla IV.17. Caracterización del Hardware	114
Tabla IV.18. Caracterización del Software.....	114
Tabla IV.19. Caracterización del software para la implantación.....	114
Tabla IV.20. Costo del Software	115
Tabla IV.21. Costo Total del Proyecto	115
Tabla IV.22. Conclusión de costos del Proyecto.....	115
Tabla IV.23. Riesgos Genéricos	122
Tabla IV.24. Códigos y efectos para categorizar el riesgo	122
Tabla IV.25. Categorización del riesgo	123
Tabla IV.26. Refinamiento de la Historia de usuario 1	129
Tabla IV.27. Refinamiento de la Historia de usuario 2	130
Tabla IV.28. Refinamiento de la Historia de usuario 3	131
Tabla IV.29. Representación de la historia de usuario1	132
Tabla IV.30. Representación de la historia de usuario2	133
Tabla IV.31. Representación de la historia de usuario3	134
Tabla IV.32 Prueba de la gestión de información general.....	158
Tabla IV.33 Prueba de la gestión de información semántica	160
Tabla IV.34 Prueba de la gestión de información semántica	163

Tabla IV.35 Valores abstraídos en base a las cuatro primeras preguntas.....	164
Tabla IV.36 Valores abstraídos en base a las cuatro últimas preguntas.....	164
Tabla IV.37 Valores esperados de las cuatro primeras preguntas.....	166
Tabla IV.38 Valores esperados de las cuatro últimas preguntas.....	166
Tabla IV.39 Chi-Cuadrado para las primeras cuatro preguntas.....	167
Tabla IV.40 Chi-Cuadrado para las últimas cuatro preguntas.....	167
Tabla IV.41. Votantes por cada pregunta.....	169

ÍNDICE DE FIGURAS

Figura II.1. Ejemplo de ontología.....	28
Figura II.2. Ejemplo de Macrotesauros	29
Figura II.3. Ejemplo de Topic Maps.....	32
Figura II.4. Componentes de un documento SGML.....	39
Figura II.5. Diagrama TAO (Topics Asociaciones Ocurrencias) de un Topic Map.....	55
Figura II.6. Ejemplo de Topic Type	56
Figura II.7. Ejemplo de Occurrence.....	57
Figura II.8. Ejemplo de Resource Data.....	57
Figura II.9. Ejemplo de Association	59
Figura II.10. Ejemplo de Association Role.....	59
Figura II.11. Ejemplo de Association en XML	60
Figura II.12. Ejemplo de TMQL.....	62
Figura II.13. Árbol con Hipervínculos o Árbol con hiperenlaces	66
Figura III.1. DTD del XTM	82
Figura III.2. Funcionamiento de XTM	83
Figura IV.1. Infraestructura tecnología del Liceo Stephen Hawking.....	91
Figura IV.2. Software Cocomo.....	108
Figura IV.3. Total costo de desarrollo según Cocomo	108

Figura IV.4. Total planes y requerimientos según Cocomo	109
Figura IV.5. Total diseño del producto según Cocomo	109
Figura IV.6. Total Programación del proyecto según Cocomo	110
Figura IV.7. Total Pruebas e Integración según Cocomo	110
Figura IV.8. Escala de Complejidad	111
Figura IV.9. Escala de ponderación	112
Figura IV.10. Fechas de entrega del proyecto	117
Figura IV.11. Diagrama Gantt	118
Figura IV.12. Representación grafica de la historia de usuario1	132
Figura IV.13. Representación gráfica de la historia de usuario2	133
Figura IV.14. Representación gráfica de la historia de usuario3	134
Figura IV.15. Diagrama del modelo conceptual	136
Figura IV.16. Diagrama de calle del modulo1	137
Figura IV.17. Diagrama de calle del modulo2	138
Figura IV.18. Diagrama de calle del modulo3	138
Figura IV.19. Diagrama UML de la segunda iteración	140
Figura IV.20. Diagrama UML de la tercera iteración	142
Figura IV.21. Diagrama UML de la cuarta iteración	143
Figura IV.22. Diagrama UML de la quinta iteración	145
Figura IV.23. Diseño relacional de la Base de Datos	154

Figura IV.24. Diagrama de Componentes	155
Figura IV.25. Estructura física de los nodos.....	156
Figura IV.26. Gráfico de resultados de las primeras cuatro preguntas de la encuesta. 168	
Figura IV.27 Gráfico de resultados de las últimas cuatro preguntas de la encuesta.....	168
Figura IV.28. Relación del número de votantes y el puntaje obtenido.....	169

INTRODUCCIÓN

El conocimiento creado a partir de una inteligencia colectiva que define las capacidades cognitivas de los individuos y los establece como parte de un gran cerebro compartido, se establece en una web social mediante una organización semántica basada en etiquetas, aunque muestra cierto significado en la organización de la información, no representa un mecanismo que pueda simular un razonamiento humano, careciendo completamente de relaciones semánticas que constituyen una estructura reticular de un pensamiento compartido.

La representación de un conocimiento social propio de una Web 2.0 nos conduce a la producción de portales semánticos, estas representan conceptos y relaciones existentes a través de enlaces, con una sintaxis específica la cual ayuda a hacer que las búsquedas se hagan más fáciles y eficientes, haciendo de esta manera que el usuario al navegar por el portal encuentre la información necesaria sin tener que salirse del tópico principal.

Este tipo de portales semánticos se los crea mediante un documento XML Topic Maps (XTM), el cual es un paradigma para la representación de la información y el conocimiento, sigue las reglas de XML, y este es multiplataforma es decir se puede consumir en cualquier plataforma por su que soporte XML, ya que la tendencia hoy en día es trabajar sobre archivos planos, por su portabilidad y usabilidad.

.

ORGANIZACIÓN DEL DOCUMENTO

El Capítulo I presenta el marco de referencia en el que se detalla el título, el problema, los objetivos, la justificación y la hipótesis de la investigación.

El Capítulo II contiene el marco teórico en el cual se realiza una presentación de los conceptos generales, introducción, escenarios de uso, representación del conocimiento, representación de la información en la web, representación del conocimiento bajo Topic Maps.

El Capítulo III comprende el estudio la representación del conocimiento bajo XML Topic Maps (XTM), definición, evolución, funcionamiento, ventajas, desventajas, y normativas para la construcción del documento XTM.

El Capítulo IV Aquí se realiza la implementación del caso práctico de la tesis es decir del Portal Semántico tomando en cuenta todo el estudio anterior.

Por último se presentan las conclusiones y recomendaciones obtenidas a partir de la presente investigación.

CAPÍTULO I

MARCO DE REFERENCIA

1.1. Título de la investigación

Desarrollo de un portal semántico bajo el paradigma XML Topic Maps (XTM):
Caso práctico Liceo Stephen Hawking.

1.2. Problema de la Investigación

En la actualidad la utilización de nuevos paradigmas para representar la información y conocimiento de forma semántica, se ha visto tan necesaria en

cuanto se refiere a facilitar la navegación del usuario, que es de vital importancia para todo el mundo, por lo cual dichos paradigmas han creado nuevas técnicas, mismas que facilitan la recuperación de información y conocimiento requerida; una de estas es por medio de XML Topic Maps (XTM).

1.2.1. Análisis

Dentro del análisis para la representación de la información y el conocimiento realizaremos el estudio del paradigma XML Topic Maps (XTM), dando a conocer su evolución, funcionamiento, ventajas normativas y el software para el desarrollo del portal semántico para el Liceo Stephen Hawking.

1.2.2 Limitación

La limitación que se presenta es la ausencia de un sistema automatizado en el Liceo Stephen Hawking.

1.3. Objetivos

1.3.1. Objetivo General

Implementar un portal semántico que represente información y conocimiento basado en XTM para el Liceo Stephen Hawking.

1.3.2. Objetivos Específicos

- Analizar la representación del conocimiento
- Analizar la representación de la información en la web.
- Estudiar la representación del conocimiento bajo Topic Maps TM
- Estudiar la representación del conocimiento en la web XML Topic Maps XTM.

- Desarrollar e implementar el portal semántico para el Liceo Stephen Hawking.

1.4. Justificación de la Investigación

1.4.1 Justificación Teórica

Con la inclusión de Topic Maps en el desarrollo web se puede mejorar aspectos como:

- Navegación ya que muestra todos los conceptos relacionados con el concepto central, siendo más intuitivo navegar en este espacio que en otro sistema de recuperación de información
- Agentes personales para mejorar la eficiencia en el acceso a la web.

Por lo tanto, de acuerdo a las mejoras que ofrece tanto la representación de la información como la del conocimiento en el desarrollo web, nos permite crear aplicaciones semánticas, que presentan mejoras a problemas como:

- **No centralizada:** problemas para garantizar integridad de la información.
- **Información Dinámica:** puede cambiar la información e incluso el conocimiento sobre esa información
- **Mucha información:** El sistema no puede pretender acaparar toda la información
- **Es abierta:** Muchos sistemas anteriores usaban la Closed World Assumption

1.4.2 Justificación Práctica

La ausencia en el ámbito académico en la web del Liceo Stephen Hawking, da la apertura a la creación de un espacio el cual dé ha conocer información relevante.

Con el siguiente tema de tesis se va a diseñar e implementar un portal semántico que permita representar información y conocimiento basado en XTM, enfocado al conocimiento e información del Liceo Stephen Hawking:

- Portal semántico para el Liceo Stephen Hawking.
 - Mapa de Tópicos Institucional
 - Mapa de tópicos Académicos

1.5. Hipótesis

La implementación del portal semántico bajo el paradigma XML Topic Maps (XTM), en el Liceo Stephen Hawking mejorará la eficiencia y navegabilidad en el acceso a la web.

CAPÍTULO II

MARCO TEORICO

2.1. Representación del conocimiento

Durante muchos años la forma de recolectar información no ha sido más que una tarea ardua tanto para los educadores como los estudiantes, para los investigadores y aprendices puesto que la manera de representar conocimiento se daba forma ambigua, ahora hay diferentes técnicas para hacer del aprendizaje algo más ordenado, dándonos a conocer ontologías, tesauros, y topic maps.

2.1.1. Ontologías

El término **ontología** se origina en el campo de la filosofía y la epistemología. Como ciencia, la Ontología es una rama de la metafísica que se ocupa del estudio de la naturaleza de la existencia, de los seres y de sus propiedades

transcendentales; en filosofía, por tanto, una ontología se considera como una explicación sistemática de la Existencia. La magnitud de esta rama del saber y sus conexiones con la epistemología (rama de la filosofía que estudia la naturaleza y las fuentes del conocimiento) hacen que el estudio de la ontología desde el punto de vista filosófico quede fuera de los alcances de nuestra investigación. Una definición de diccionario típica del término ontología la identifica con "la rama de la metafísica que estudia la naturaleza de la existencia".

Una definición con respecto a conceptos dice que una **ontología** es una descripción formal de los *conceptos* y las *relaciones* entre conceptos.

Por lo tanto **Ontología** es una *descripción formal de conceptos* en el *dominio* de un discurso. Ontologías son teorías que especifican un vocabulario relativo a un cierto dominio. Este vocabulario define entidades, clases, propiedades, predicados y funciones y, las relaciones entre estos componentes. Las ontologías toman un papel clave en la resolución de interoperabilidad semántica entre sistemas de información y su uso dentro del contexto.

En términos prácticos, el desarrollo de una ontología incluye:

- ✓ Definir *clases* en la ontología
- ✓ Colocar las clases en un *jerarquía de taxonomías* (subclase-superclase)
- ✓ Definir [*slots* –atributos–] y describir los valores permitidos para esos [*slots*]
- ✓ Rellenar los valores de los slots con ejemplos.

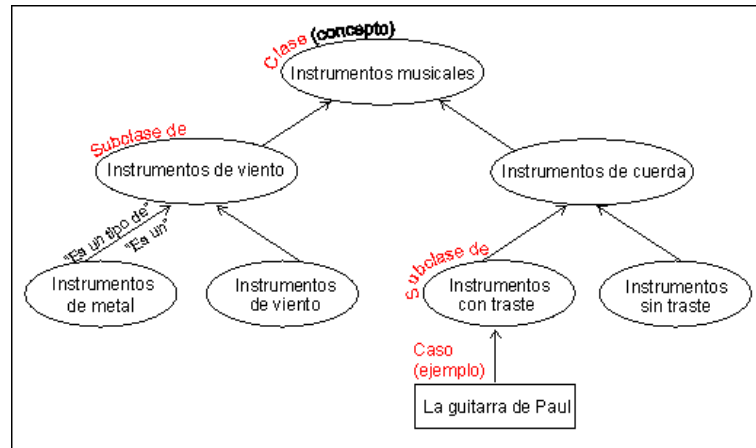


Figura II.1. Ejemplo de ontología

Partiendo de este significado, el término ontología es usado en el ámbito de la *organización del conocimiento* para referirse a un conjunto de conceptos organizados jerárquicamente, representados en algún sistema informático cuya utilidad es la de servir de soporte a diversas aplicaciones que requieren de conocimiento específico sobre la materia que la ontología representa.

2.1.2. Tesauros

Si definimos un tesoro por su función, es un instrumento de control terminológico utilizado para trasponer a un lenguaje más estricto el idioma natural de los documentos. Por su estructura, es un vocabulario controlado y dinámico de términos que tienen entre ellos relaciones semánticas y genéricas, y que se aplica a un dominio concreto del conocimiento. El componente esencial de los tesauros son los descriptores, palabras o expresiones corrientes escogidas por el constructor del tesoro para designar los conceptos representativos del documento.

ELEMENTOS DE UN TESAURO

- **Unidades lexicales.** Estas pueden admitir las siguientes categorías: grupos de descriptores, que agrupan los términos de indización bien por campos (temas) o por clases de términos o "facetas"

- **Los descriptores.** Son palabras o expresiones del lenguaje corriente retenidas por el constructor del tesoro para escoger los conceptos representativos del documento y las preguntas utilizadas por el indizador.
- **No descriptores.** Son sinónimos o cuasi-sinónimos de los descriptores o términos que designan en el lenguaje de uso conceptos afines a los que cubren los descriptores. Los no descriptores no pueden ser utilizados para la indización del documento o de las preguntas, pero cada uno de ellos reenvía a uno o dos descriptores para representar los conceptos correspondientes.

TIPOLOGIA

Pueden concebirse desde dos enfoques:

- *Por facetas*, que según la norma ISO 2788 sobre la construcción de tesauros monolingües, es una noción más abstracta que el concepto de campo, lo cual requiere una disciplina mental más rigurosa para el compilador. Los documentos se agrupan por clases o por puntos de vista, según un tipo básico de conceptos.
- *Por campos*, que pone en relación los conceptos con los cuales generalmente se asocia un concepto.

OECD Macrothesaurus - DEMANDA

ID: 1398
 PC: DEMANDA
 TM: OFERTA Y DEMANDA
 TE: DEMANDA DE CONSUMO
 TE: DEMANDA DE ENERGIA
 TE: DEMANDA DE MANO DE OBRA
 TE: NECESIDADES BASICAS
 TE: NECESIDADES FINANCIERAS
 TR: ADMINISTRACION DE LA DEMANDA
 FA: 09.01.01

[en | es | fr] [INDEX] [A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U | V | W | Y | Z]

© OECD: HTML-Version by [gaj](#)

Figura II.2. Ejemplo de Macrotesauros

2.1.3. Topic Maps

Un mapa conceptual es una técnica para representar y organizar el conocimiento empleando conceptos y frases de enlace entre estos conceptos. Un mapa conceptual, según la propuesta de Novak¹, es un conjunto de proposiciones sobre un determinado tema ordenadas en forma de árbol. En los nodos de este árbol se colocan los conceptos y en las conexiones entre estos nodos irían situadas frases de enlace expresando la relación entre los conceptos conectados. Normalmente los conceptos son la parte sustantiva de las oraciones y los verbos o preposiciones constituyen las frases de enlace. Siempre hay un concepto inicial (raíz) a partir del cual se desarrolla el árbol de relaciones. Los conceptos se suelen representar dentro de cuadrados u óvalos y las frases de enlace actúan de etiquetas de las líneas que unen dos o más de estos cuadrados de conceptos. Los mapas conceptuales son instrumentos con orígenes, funciones y objetivos distintos. Por un lado, los mapas conceptuales tienen su origen en un contexto académico, su función inicial fue la mejora del aprendizaje y su principal objetivo es representar el conocimiento en forma gráfica. En cambio, *Topic Maps* es una norma internacional impulsada en sus orígenes por los productores de libros electrónicos con el objetivo de almacenar y procesar información en un formato estandarizado los puntos en común, que no son pocos. Como se mencionó, ambas realidades obedecen a un mismo modelo teórico en cuanto a la estructura de la información: grafo dirigido. No se detallará la descripción de las peculiaridades de este tipo de estructura, simplemente destacar que en ambos casos, hay nodos (conceptos, *topics*) relacionados mediante frases de conexión formando una red de iguales características. Esta coincidencia hace posible que *Topic Maps* pueda ser un formato de almacenamiento de los mapas conceptuales, o formulado de otra manera, que usando la norma *Topic Maps* se pueda crear de forma muy natural un mapa conceptual.

La visualización de la información contenida en un *Topic Maps* en forma de mapa conceptual es solo una de las posibles formas de mostrar su contenido. El

¹ Gerardo González García[1] entrevista a Novak que es un Experimentado Investigador Científico, desarrolló los Mapas Conceptuales, como ahora se los conoce

formato *Topic Maps* y el modelo conceptual que implica puede aplicarse a otros fines, como por ejemplo la recuperación de la información en base a inferencias automáticas.

Está claro que lo que proponen estas teorías y los autores es, entre otras cosas, que los mapas conceptuales

- Deben ser simples, y mostrar claramente las relaciones entre conceptos y/o proposiciones.
- Van de lo general a lo específico, las ideas más generales o inclusivas, ocupan el ápice o parte superior de la estructura y las más específicas y los ejemplos la parte inferior. Aún cuando muchos autores abogan porque estos no tienen que ser necesariamente simétricos.
- Deben ser vistosos, mientras más visual se haga el mapa, la cantidad de materia que se logra memorizar aumenta y se acrecienta la duración de esa memorización, ya que se desarrolla la percepción, beneficiando con la actividad de visualización a estudiantes con problemas de la atención.
- Los conceptos, que nunca se repiten, van dentro de óvalos y la palabras enlace se ubican cerca de las líneas de relación.
- Resulta conveniente escribir los conceptos con letra mayúscula y las palabras de enlace en minúscula, pudiendo ser distintas a las utilizadas en el texto, siempre y cuando se mantenga el significado de la proposición.
- Para las palabras enlace se pueden utilizar verbos, preposiciones, conjunciones, u otro tipo de nexo conceptual, las palabras enlace le dan sentido al mapa hasta para personas que no conozcan mucho del tema.
- Si la idea principal puede ser dividida en dos o más conceptos iguales estos conceptos deben ir en la misma línea o altura.
- Un mapa conceptual es una forma breve de representar información.

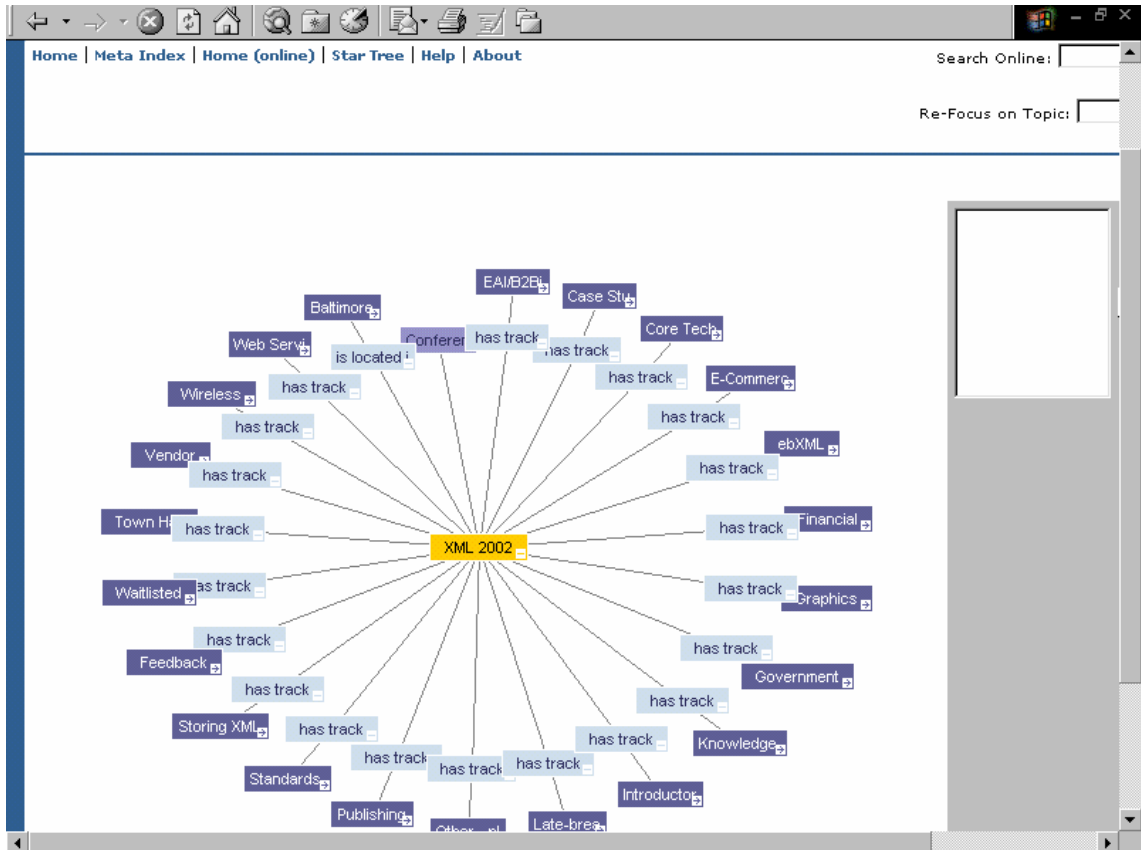


Figura II.3. Ejemplo de Topic Maps

Resumen

CARACTERÍSTICAS DE LAS DIFERENTES FORMAS DE REPRESENTACIÓN DEL CONOCIMIENTO

Mapa conceptual	Topic Maps	Tesauros	Ontologías
No existe norma	ISO 13250:2000	ISO 2788:1986 (o equivalente UNE 50-106-90)	
Control del vocabulario. El mapa conceptual no realiza control de vocabulario a priori. En el proceso de	El vocabulario de un <i>topic map</i> viene dado por el documento, los documentos que sirven de base para	Los descriptores, y no descriptores, se marcan a priori.	El vocabulario posee entidades, clases, propiedades, predicados y funciones y, las

<p>aprendizaje <i>profesor</i> y <i>alumno</i> marcan los conceptos y los nodos.</p>	<p>generarlo, y son los términos que van marcando de manera dinámica cuáles son los conceptos subyacentes a cada <i>Topic</i></p>		<p>relaciones entre estos componentes</p>
<p>Dominios. El mapa conceptual se orienta a dominios, pero la validez de cada nodo es algo subjetivo que establece la persona que elabora el mapa conceptual.</p>	<p><i>Theme/Scope</i>, marca el límite de la validez de las asignaciones de cada <i>topic</i>, límite que va variando con los sucesivos <i>themes</i> en los que van apareciendo los <i>topics</i>, que también pueden ser infinitos.</p>	<p>Relación de Jerarquía (Generalización y enumeración). Organización por campos y organización por indicadores clasificatorios (UNE 50-106-90: 9.3.2.1 y 9.3.2.2)</p>	<p>Todos y cada uno de los términos deben estar asignados a un concepto determinado. Cada uno de los conceptos, por su parte, debe formar parte de una estructura bien definida y debe ser posible especificar diversos tipos de relaciones entre ellos</p>
<p>Tipos de nodos. No existe una tipología de nodos en el mapa conceptual.</p>	<p><i>Topic types</i>, que representan una relación del tipo clase-instancia, van variando en función del tipo de información (<i>Theme/Scope</i>) que</p>	<p>Descriptores y no-descriptores de un dominio.</p>	<p>Se definen clase, jerarquía de taxonomías y slots – atributos</p>

	estemos tratando.		
Relaciones. Son arcos etiquetados (verbo o preposición, verbo o nombre, verbo o conector lógico) entre conceptos, nodos.	Las relaciones son etiquetadas sólo por verbos. Son estructuras del tipo [<i>topic</i>] - [verbo] - [<i>topic</i>]. El documento o documentos que sirven de base para generar un <i>topic map</i> son los que van marcando de manera interactiva cuáles son los conceptos subyacentes a cada <i>topic</i> .	No existe un conector entre los términos relacionados en un tesauro.	Se especifica diversos tipos de relaciones siendo estos por ejemplo: “Es un”, “Es un tipo de”.
Relaciones. Los mapas conceptuales suelen establecer relaciones del tipo: ‘ <i>being</i> ’, ‘ <i>having</i> ’, ‘ <i>using</i> ’, ‘ <i>causing</i> ’, ‘ <i>showing</i> ’, ‘ <i>including</i> ’, ‘ <i>similarity</i> ’, pero no existe un acuerdo o norma que los fije.	<i>Association types</i> . Los <i>Topic Maps</i> pueden tener un número de relaciones potencialmente infinito.	Relación de Jerarquía Relación de Equivalencia Relación de Asociación	Axiomas para relaciones entre conceptos
El conjunto ‘nodo-enlace-nodo’ No puede ser	Una <i>association type</i> o, incluso, un <i>topic type</i> pueden ser <i>topic</i>	Las relaciones que se establecen en un tesauro no suelen ser	Los conceptos deben estar bien estructurados y

considerado un nodo.	puede ser una <i>association</i> .	consideradas descriptores.	relacionados entre ellos
Equivalentes ortográficos. No existe esta concreción en el mapa conceptual (trabaja con conceptos con independencia de los términos en los que se exprese).	<i>Topic names</i> , que son las variaciones ortográficas de un mismo término, apodos, acrónimos	Relación de Equivalencia, sólo a través de este tipo de relación es posible contemplar algunos casos similares.	Se considera como el repositorio de conceptos que establecen conexiones entre los símbolos de una lengua y sus referentes en el campo o subcampo que se estudia
Los nodos no dependen de los documentos en los que puedan aparecer.	<i>Occurrences</i> , casos relevantes del <i>topic</i> que están fuera del propio <i>topic map</i> que se está tratando.	<i>Scope Notes</i> Origen del término Definición del término	Los campos y subcampos son independientes de sus diferentes relaciones
El mapa conceptual establece conceptos relevantes en el dominio, pero no términos.	<i>Public subject</i> , que pueden ser tanto los <i>topics</i> que representan al mismo <i>subject</i> , como las variaciones idiomáticas de un <i>topic</i> , quedando todos al mismo nivel.	No descriptor. Nivel diferente. Tesoro multilingüe.	Los conceptos relevantes en un dominio específico pueden crear variaciones idiomáticas
Dominios semántico - conceptuales sólo válidos en los límites	<i>Facets</i> , expresadas mediante pares atributo valor, y que	Facetas (predefinidas en el tesoro)	Representación mediante el campo y

de cada mapa conceptual	pertenecen a la descripción formal del <i>topic</i> y no a la semántico-conceptual.		subcampo establecido por las reglas ontológicas
-------------------------	---	--	---

Tabla II.1. Características de las diferentes formas de Representación del Conocimiento

2.2. Representación de la información en la web

Ahora si bien vimos que la representación de conocimiento se dio en ámbitos diferentes logrando facilitar el aprendizaje y la recolección de datos para consultas, sin embargo esta tarea sigue siendo un tanto difícil, ya que la tecnología va avanzando, creando de esta manera más espacio para información presentada en diferentes formatos, para esto se crean estándares, los cuales mejorarán la estructuración y representación de la información en la web, mostrándonos entre otros SGML como uno de los primeros, HTML, XML y XHTML como evolución al paso del tiempo.

2.2.1. SGML (Standard Generalized Markup Language)

- Creado en 1969 en IBM por Charles Goldbarb
- En 1986 se convierte en norma ISO 8879-1986
- Se basa en el ‘marcado descriptivo’ o ‘generalizado’, que indica cómo se deben intercalar marcas en un documento para diferenciar sus componentes estructurales.
- Las marcas no señalan cómo se debe presentar el documento en pantalla ni cómo se tiene que formatear el documento al imprimirlo

SGML no propone un conjunto de marcas predefinidas, sino:

- La sintaxis que debe utilizarse para definir un conjunto de marcas aplicables a los documentos de un mismo tipo (aplicación)

- La forma en la que se debe intercalar estas marcas en los documentos
- Los juegos de caracteres que se pueden utilizar en los documentos (ISO 646 e ISO 10646)
- Los elementos que pueden aparecer en un tipo de documento, sus características y el orden en el que deben escribirse se definen en un documento aparte llamado DTD (Document type definition)
- Los documentos se consideran ‘instancias’ de un tipo de documento específico que define su estructura válida
- Un documento SGML siempre debe cumplir las restricciones que se indican en su DTD
- Un documento SGML debe incluir una referencia a la DTD a partir de la cual se ha definido
- SGML tiene capacidad hipertexto basadas en la norma HyTime
- Los enlaces SGML unen un elemento origen con un elemento destino
- El elemento origen tendrá un atributo de tipo IDREF (por convención se llama link o xref a este atributo)
- El elemento destino tendrá un atributo de tipo ID (por convención se llama target)

Funcionamiento de SGML

La función del etiquetado es aportar información que refleje la estructura jerárquica de un documento, de forma que ayude al lector u ordenador a procesar su contenido. Existen dos tipos de etiquetado:

1. El **etiquetado procedimental** : técnica mediante la cual un operario utiliza instrucciones crípticas y dependientes del funcionamiento de un sistema determinado para que ejecute una acción, como por ejemplo activar la tipografía, poner en negrita, centrar, etc.
2. El **etiquetado descriptivo**, que identifica los elementos estructurales de un documento, determinando su estructura lógica.

El lenguaje de etiquetado SGML permite distinguir entre el contenido (**datos**) y el **etiquetado**.

Un documento SGML consta de 3 partes:

- Declaración SGML
- Declaración de tipo de documento (DTD)
- Instancia de Documento

Muestra de documento

La Muestra de Documento contiene el marcado o etiquetado y el contenido del texto, es decir, contiene el documento en sí que incluye tanto el texto como el marcado.

El contenido del documento tiene objetos SGML que siguen la estructura del árbol definido en la DTD. Los objetos SGML son fundamentalmente dos:

- Elementos (insertados como etiquetas y sus atributos locales)
- Entidades (para caracteres especiales, texto almacenado y archivos externos).

En la siguiente imagen pueden distinguir los diferentes componentes de un documento SGML

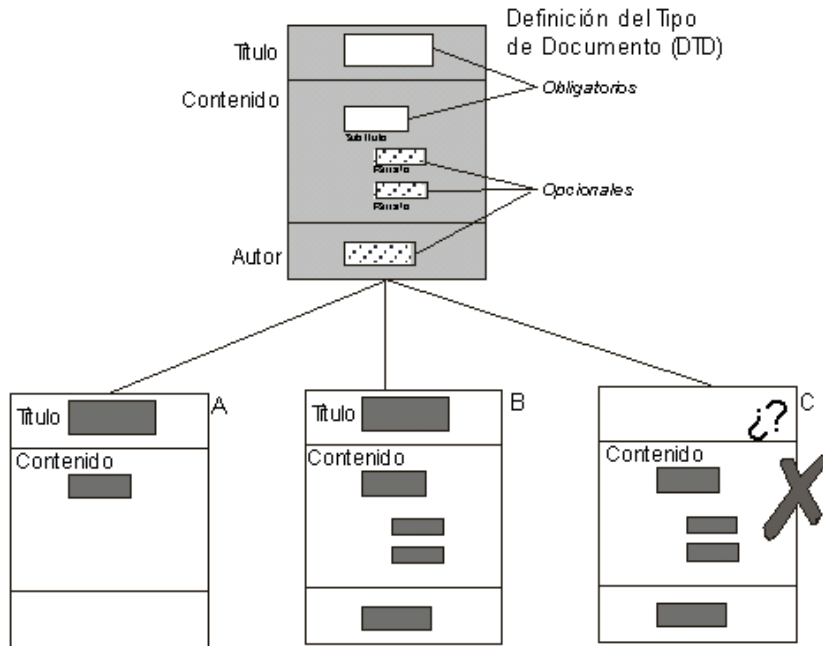


Figura II.4. Componentes de un documento SGML

Características de un documento electrónico en SGML

El material que constituye un documento se puede distribuir en diferentes archivos, que pueden estar almacenados en distintos ordenadores.

En SGML cada uno de ellos recibe el nombre de entidad; éstas pueden tener cualquier tamaño, haber sido creado por cualquier programa de software o estar guardadas en cualquier ordenador, y se conciben como objetos independientes.

Las entidades pueden estar compartidas por distintos documentos.

Un documento estará definido en función de la estructura de las entidades que lo conforman.

Procesamiento de un documento SGML

Se trata de una validación automática realizada por una máquina.

La función del parser² es leer el documento SGML y separar los datos del etiquetado. Detecta cuándo el etiquetado ha sido minimizado y en tal caso lo expande. Si el contenido incluye referencias a una hoja de cálculo electrónica de un capítulo concreto, y el gráfico del organigrama de otro capítulo, dará las instrucciones pertinentes al sistema sobre cómo encontrar dichas entidades.

Si el gráfico se halla en alguna notación de contenido de datos especial, generada por un programa de diseño de gráficos, el parser lo dispondrá todo para introducir la imagen.

Si su contenido incluye instrucciones especiales para el sistema de edición en su propio lenguaje interno (instrucciones de procesamiento) éstas pasarán directamente a la aplicación.

Si se ha utilizado el componente de sección marcada en SGML y ha indicado que algunas partes de su documento no han de aparecer en la versión editada, el parser sabrá que no tiene que enviarlas.

Si se está utilizando el componente de declaración de comentarios SGML para enviar y recibir notas y mensajes entre los escritores y los editores, el parser también sabrá que no ha de enviarlos a la aplicación receptora.

Todo esto se realiza sin que resulte visible al ojo humano.

Validación del documento SGML

Antes de generar cualquier tipo de documento a partir de SGML, es necesario que se trate de un documento válido, es decir que se respete la sintaxis de DocBook³. Para ello basta usar el comando C-c C-v Enter, en caso de que no se haya salvado aparecerá un mensaje en el buffer pidiendo hacerlo, para lo cual se debe escribir y.

² ALEUSA [2], parser es un programa o módulo encargado del análisis de sintaxis de un texto.

³ Wikipedia[3] **DocBook** es un aplicación del estándar SGML/XML e incluye una DTD propia

A continuación se dividirá la ventana y aparecerá un mensaje que indica el resultado de la validación correcta.

2.2.2. Tecnología HTML (Hypertext Markup Language)

- HTML es una aplicación SGML para codificar documentos y distribuirlos en el World Wide Web
- HTML define un conjunto limitado de marcas que se pueden intercalar en los documentos
- Comparte las ventajas de SGML: multiplataforma, fácilmente procesable, etc.
- Ha alcanzado un gran éxito, hasta el punto de llegar a igualarse edición electrónica con edición HTML
- La simplicidad de HTML acarrea limitaciones:
 - ✓ Falta de fórmulas de compresión asociadas al formato (problema ancho de banda)
 - ✓ Posibilidades de formateo muy limitadas
 - ✓ Falta de mecanismos de acceso: búsqueda texto completo, tablas de contenidos, etc.
 - ✓ Escasa capacidad expresiva del lenguaje
 - ✓ Disponemos de browsers y tecnologías (Java, lenguajes de script, etc.) muy potentes cuya capacidad está siendo infrutilizada por las limitaciones del formato HTML
- La evolución de HTML ha estado condicionada por la presión ejercida por los fabricantes de los browsers más utilizados
- dHTML hace referencia a unas características soportadas por la versión 4 de los browsers de Netscape y Microsoft
- Algunas de estas características se tomaron del borrador disponible para la versión 4 del formato HTML
- dHTML permite:

- ✓ Control sobre los elementos que conforman una página HTML mediante el DOM (Document Object Model)
 - ✓ Posibilidad de formatear los elementos de la página
 - ✓ Posibilidad de modificar el formato de los elementos en respuesta a acciones realizadas por el usuario
 - ✓ Control de la posición de un elemento u objeto dentro de la página
 - ✓ Capacidad de cambiar el contenido de la página una vez ésta ha sido descargada por el navegador
- Problemas de dHTML
 - ✓ Orientado a la presentación de los documentos y a aspectos ‘visuales’
 - ✓ No representa el contenido semántico ni la lógica de la información que contiene el documento
 - ✓ Las ‘implementaciones’ de Microsoft y de Netscape no son idénticas: las mismas hojas de estilo son interpretadas de forma distinta por los navegadores, problemas en soporte a lenguajes de script, controles ActiveX, etc.

HTML: Origen y objetivos

- HTML 2.0 (noviembre 1995)
- HTML +, HTML 3.0
- HTML 3.2 (enero 1997)
- HTML 4.0 (diciembre 1997, abril 1998)
- Mejoras: hojas de estilo, cuadros, objetos embebidos, textos de derecha a izquierda y direcciones mixtas, mejores tablas, enlaces a formularios y mejora el acceso para personas con discapacidades
- Adopción del Estándar Internacional ISO/IEC 10646 (lenguajes universales)

HTML 4.0: Características

- Publicar documentos online con encabezados, textos, tablas, listas, fotos, etc.
- Recuperar información online vía links de hipertextos
- Diseñar formularios para transacciones con servicios remotos, para buscar información, hacer reservaciones, ordenar productos, etc.
- Incluir spread-sheets, video clips, sound clips y otras aplicaciones relacionadas con los documentos
- Distingue mejor entre estructura y presentación
- Mejores formularios, claves de acceso, agrupamiento de controles, agrupa las opciones de selección
- Posibilidad de marcar (tags) la descripción de un texto de un objeto incluido (con el elemento objeto)
- Mecanismo (map) para que un autor pueda integrar imágenes con links de texto
- El requerimiento que alterna texto acompañando imágenes incluidas con el elemento `` y los mapas de imágenes incluyendo el elemento *area*
- Soporte para atributos título `<title>` y lenguaje `<lang>` en todos los elementos, abreviatura `<abbr>` y acrónimo `<acronym>`
- Un rango mayor de objetivos (target media) para usar con hojas de estilo
- Mejores tablas incluyendo nombres, grupos de columnas y mecanismos no visuales de interpretación

HTML 4.0: Tablas

- Mayor control sobre la estructura y diseño (layout)
- Despliega los datos de la tabla de forma incremental

HTML 4.0: Documentos Compuestos

Mecanismo estándar para embeber objetos media genéricos y aplicaciones.
 Provee un mecanismo para incluir:

- Imágenes, video, sonido, elementos matemáticos, y aplicaciones especiales
- Permite especificar la jerarquía de cargado de contenido alternativo para visualizadores que no soporten un cargado específico

HTML 4.0: Hojas de estilo

- Simplifican el sistema de marcas de HTML
- Les otorga, tanto al autor como al usuario, el control sobre la presentación de los documentos (información de la fuente, alineamiento, color)
- Pueden estar embebidas en un documento HTML o en archivos externos
- El mecanismo para asociar una hoja de estilo al documento es independiente del lenguaje de la hoja de estilo

HTML 4.0: Usos de scripts

- Páginas Web Dinámicas
- Usan HTML para construir aplicaciones en red
- El mecanismo para incluir scripts a un documento HTML, es independiente del lenguaje del script

HTML 4.0: Impresión

Posibilita imprimir más allá de la página en pantalla creando relaciones entre documentos:

- Elemento link de HTML
- RDF: Resource Description Language

2.2.3. Tecnología XML (eXtensible Markup Language)

- Su desarrollo comienza en septiembre de 1996 dirigido por el W3C y con la participación de importantes empresas: Microsoft, IBM, Sun, Novell, ArborText, H-P etc.

- El propósito es:
 - ✓ Diseñar un lenguaje de marcas optimizado para el WWW
 - ✓ Unir la simplicidad de HTML con la capacidad expresiva de SGML
- Versión 1.0 ratificada en diciembre de 1997
- Áreas de aplicación:
 - ✓ Representación y distribución de documentos e información textual
 - ✓ Intercambio de datos e información estructurada a través de Internet y el WWW
 - ✓ Integración de datos procedentes de fuentes heterogéneas
- Elimina la ‘barrera’ entre información estructurada e información textual

Origen y objetivos de XML

Es una actividad de la W3C para viabilizar la utilización de SGML en la WWW. Desarrollado entre 1996 y 1998 por un conjunto de expertos en lenguajes de marcas provenientes de la industria y la academia.

- Eliminará la complejidad de SGML (difícil y costosa elaboración de programas y su utilización)
- Posibilitará utilizar el SGML genérico de un modo tan sencillo como hoy es utilizado el HTML
- XML será interoperativo con HTML y SGML

Características de XML

- Reglas fáciles de seguir para crear un lenguaje de marcas determinado desde cero.
- Las marcas no tienen un significado predeterminado
- Cualquiera puede entenderlas porque consisten en texto común

- Transmite contenido y estructura
- No transmite presentación ni comportamiento. Para que lo haga se le debe asociar mecanismos adicionales: Hojas de estilo (XSL o CSS)

Conceptos acerca del documento XML

Existen tres elementos claramente definidos:

- Contenido
- Estructura
- Presentación

XML y las DTDs

- SGML utiliza DTDs para la descripción de los elementos disponibles en un tipo de documento dado
- El diseño y construcción de una DTD no es trivial
- XML ha sido diseñado para ser utilizado con o sin DTDs
- Documento sin DTD: define sus propias marcas a través de la existencia y localización de elementos en el momento en que es creado

Un visualizador XML necesita entender la estructura de lo que está leyendo, esto introduce el concepto de:

- Documento-bien-formado
- Documento-válido (tiene una DTD asociada)
- XML

Crearé aplicaciones en la Web centradas en los datos:

- Intercambio entre bases de datos
- Agentes inteligentes podrán manipular el contenido de la información de acuerdo a las necesidades de los clientes
- Gestión de las colecciones de documentos

- Cambiará las publicaciones en la Web y, posteriormente también a las publicaciones en general
- SGML y XML especifican el contenido y la estructura de un documento de forma tal que permiten generar distintas presentaciones de acuerdo a la necesidad.
- Los distintos sistemas operativos serán transparentes para el usuario así como los idiomas (transculturalismo)

Hacia el Formato Universal

- El futuro: un formato universal XML + XSL
- La combinación XML+XSL puede remplazar a todos los procesadores de texto y formatos existentes
- Un único formato para publicaciones impresas y electrónicas
- Un único formato para diferentes tipos de productos
- Un único formato para todos los idiomas

Implicancias económicas y políticas de la asociación XML + XSL

- Los usuarios no estarán más atados a un tipo comercial de software
- Cambiarán las relaciones entre vendedores de software y clientes
- Será el final de la dominación de unas pocas empresas gigantescas y de unos pocos países poderosos

Ventajas de utilizar XML en las aplicaciones Web.

Entre las muchas que existen podemos destacar tres:

- Sencillez
- Variedad de estructuras de datos
- Excelente tratamiento de caracteres internacionales.

Sencillez

La primera y más importante ventaja del XML se refiere a su sencillez, en especial si se lo compara con los formatos binarios.

El XML es un formato basado en caracteres y por tanto comprensible para los seres humanos. Además los documentos XML pueden leerse fácilmente, crearse y modificarse por medio de las herramientas que utilizamos normalmente, como editores de texto. Todo esto hace que la compresión y el análisis de documentos XML resulten mucho más sencillos que los escritos en formato binario.

Otro ejemplo de la sencillez de XML tiene que ver con su habilidad para representar datos estructuras en forma de árbol con todas las ventajas que esto comporta. Existen otras sintaxis estándar, como por ejemplo la ANS.1 (Abstract Syntax Notation 1) que nos permiten representar datos estructurados en forma de árbol. Pero dicho estándar resulta bastante complicado de entender y generar. Se pueden utilizar herramientas de generación automática para facilitar la tarea, pero las herramientas adecuadas suelen ser bastante costosas. Por el contrario el XML posee la misma habilidad para representar datos estructurados en forma de árbol y su compresión y manejo son más sencillos.

Si algo puede decirse del intrincado mundo de Internet es que la "sencillez gana y la eficacia pierde". En Internet la regla de oro es la apertura, es decir, accesibilidad y disponibilidad para todos. Incluso si se trata de una tecnología totalmente revolucionaria, sólo logrará imponerse si cuenta con el apoyo de la mayoría de la población afectada. Una tecnología menos eficaz pero abierta a todo el mundo y fácilmente comprensible tendrá más probabilidad de imponerse en Internet que otra mejor pero de acceso restringido. Por esto ha triunfado el HTML y por esta razón está triunfando el XML: es sencillo.

Variedad de estructuras de datos

Aunque sencillo, XML tiene la potencia suficiente para expresar estructuras complejas de datos. Para muchas aplicaciones, una estructura en forma de árbol es lo suficientemente general y potente como para expresar datos con un cierto nivel de complejidad. De hecho, supone un buen equilibrio entre el nivel de expresividad y sencillez. Incluso estructuras de datos tan complicadas como gráficos (VML, SVG) pueden llegar a representarse en un árbol. Por todo esto, XML permite expresar estructuras complejas de datos que satisfacen las exigencias de casi todas las aplicaciones.

Tratamiento de caracteres internacionales

Una de las grandes ventajas del XML que no debe subestimarse es su capacidad para gestionar conjuntos de caracteres internacionales. Incluso si se está diseñando un documento muy simple, ésta sola característica basta para decantarse por XML.

Hoy en día, los negocios se realizan a escala mundial. Esto es especialmente cierto cuando se trata de aplicaciones Web, puesto que Internet se ha encargado de borrar las fronteras nacionales. Resulta muy común que las transacciones comerciales contengan, por ejemplo, nombres de calles en chino o nombres propios de origen árabe. La recomendación 1.0 de XML está definida de acuerdo con el conjunto de caracteres ISO-10646(Unicode), por lo que virtualmente todos los caracteres que actualmente se utilizan en el mundo son caracteres oficiales.

2.2.4. Tecnología XHTML (eXtensible Hypertext Markup Language)

XHTML es una familia de módulos y tipos de documentos que reproduce, engloba y extiende HTML 4.0. Los tipos de documentos de la familia XHTML están basados en XML, y diseñados fundamentalmente para trabajar en conjunto con aplicaciones de usuario basados en XML.

XHTML 1.0 (esta especificación) es el primer tipo de documento de la familia XHTML. Es una reformulación de las tres definiciones de tipo de documento HTML 4.0 como aplicaciones de XML 1.0. Su finalidad es ser usado como lenguaje de contenidos que es a la vez conforme a XML y, si se siguen algunas sencillas directrices, funciona en aplicaciones de usuario conformes con HTML 4.0. Los desarrolladores que migren aplicaciones hacia XHTML 1.0 apreciarán las siguientes mejoras:

- Los documentos XHTML son conformes a XML. Como tales, son fácilmente visualizados, editados y validados con herramientas XML estándar.

- Los documentos XHTML pueden escribirse para que funcionen igual o mejor que lo hacían antes tanto en las aplicaciones de usuario conformes a HTML 4.0 como en las nuevas aplicaciones conformes a XHTML 1.0.
- Los documentos XHTML pueden usar aplicaciones (e.g. scripts y applets) que se basen ya sea en el Modelo del Objeto Documento (DOM) de HTML o XML.
- A medida que la familia XHTML evolucione, los documentos conformes a XHTML 1.0 estarán más preparados para interactuar dentro de y entre distintos entornos XHTML.

Reformulación del HTML 4.0 en XML 1.0

XHTML es la base para una futura familia de tipos de documentos que son extensión y subconjunto de HTML.

Para qué se necesita XHTML

- XHTML es diseñado para ser extensible
- XHTML es diseñado para ser portable

Características

- Es un reemplazo del HTML tradicional
- Es una versión más estricta y limpia del HTML
- Se define como una aplicación XML
- Es una recomendación del W3C

EL DOCTYPE

Todos los documentos XHTML válidos deben llevar un elemento llamado DOCTYPE, el cual no es parte del documento en sí, sino que define el tipo de DTD (Document Type Definition o Definición de tipo de documento) a emplear en los documentos, es obligatorio y puede ser uno de estos tres:

- XHTML 1.0 Strict: Se usa cuando se desea utilizar al 100% XHTML, su nombre lo dice bien claro, es XHTML estricto, la declaración del mismo es como sigue: `<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">`
- HTML 1.0 Transitional: Es el más usado ya que permite manejar elementos de XHTML y HTML 4.01, además de que se debe usar cuando el navegador no soporta correctamente CSS, su declaración es la que sigue: `<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">`
- XHTML 1.0 Frameset: Se debe usar cuando se manejan frames, su declaración es la siguiente: `<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Frameset//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-frameset.dtd">`

CARACTERÍSTICAS DE LOS DIFERENTES TIPOS DE REPRESENTACIÓN DE LA INFORMACIÓN EN LA WEB

SGML	HTML	XML	XHTML
Lenguaje Original	Aplicación de SGML	Perfil de SGML	Aplicación de XML y HTML
Se basa en el 'marcado descriptivo' o 'generalizado', que indica cómo se deben intercalar marcas en un documento para diferenciar sus componentes estructurales.	Conjunto limitado de etiquetas y un único tipo de documento	No especifica etiquetas, sino cómo deben definirse conjuntos de etiquetas aplicables a un tipo de documento	Se dice que es de código limpio, porque separa el contenido del diseño

Los enlaces SGML unen un elemento origen con un elemento destino	Modelo de hiperenlaces simple (unidireccionales y fijos)	Modelo de hiperenlaces complejo (múltiples destinos, fijos y relativos, etc.)	Modelo de hiperenlaces complejo (múltiples destinos, fijos y relativos, etc.)
La función del parser es leer el documento SGML y separar los datos del etiquetado. Se trata de una validación automática realizada por una máquina.	Escasa capacidad de procesamiento, el browser es un mero visor de páginas	Gran capacidad para procesar documentos, el browser es una plataforma para el desarrollo de aplicaciones	Se puede procesar documentos de manera que el browser es una plataforma para desarrollo de aplicaciones
Gran complejidad que dificulta su tratamiento e implementación, utiliza browsers específicos	El problema de la 'no compatibilidad' y las diferencias entre browsers ha alcanzado un punto en el que la solución es difícil	Fin de la guerra de los navegadores y etiquetas propietarias	Compatibilidad con aplicaciones de usuario pasadas y futuras
Validez requerida: todos los documentos deben cumplir las restricciones de su DTD	No es necesario que se cumplan las reglas de validación de documento es decir de apertura y cierre	Un documento no debe ser validado obligatoriamente, basta con que esté 'bien formado'	Además de estar bien formado se deben cumplir con las especificaciones requeridas por los DTD
Su complejidad	Siendo que la	La simplicidad de	Su coste es bajo ya

hace que las aplicaciones informáticas para procesar SGML sean muy costosas	solución es difícil y su “no compatibilidad” la solución tiene a ser costosa	XML hace más fácil el desarrollo de aplicaciones de bajo coste	que en la actualidad las aplicaciones se basan en XML que es portable y fácil
Escaso impacto al margen de sectores muy específicos: editoriales, doc.técnica	Poco soporte de la industria informática	Amplio soporte de la industria informática y más áreas de aplicación	La industria informática y muchas más áreas brindan gran apoyo puesto que está basado en XML
No hay una compatibilidad con HTML definida	Hay compatibilidad reducida con SGML y XML	Compatibilidad e integración con HTML (soporte a CSS y data-islands)	Es compatible con SGML, HTML y XML además de soportar diferentes hojas de estilo
Formateo y estilos relativamente complejos (norma DSSSL)	Formateo fácil de aplicar y estilo complicado	Formateo y estilos fáciles de aplicar mediante: CSS y XSL	Formateo y estilos fáciles de aplicar basados en las reglas de XML y CSS

Tabla II.2. Características de los diferentes tipos de representación de la información en la web

2.3. Representación del conocimiento bajo Topic Maps TM

Como un gran paso en cuanto a representación del conocimiento, ha evolucionado ha organización, y representación de la información en la web, y a pesar de los estándares establecidos para mejorar esta tarea, se presenta una técnica para representación del conocimiento basado en el estudio de Topic Maps el cual hace la recuperación y organización de la información y el conocimiento más intuitivo,

eficiente, estructurando de mejor manera la información y creando enlaces semánticos que facilitarían el aprendizaje.

2.3.1. Evolución de TM

El Topic Map tiene su origen en el grupo de Davenport, un foro destinado a productores de libros electrónicos que surgió a principios de la década de los 90.

Los proveedores se encontraban bajo la presión de los clientes para mejorar la consistencia de su documentación impresa. Eran conscientes de uso inconsistente de términos en la documentación que acompañaba a sus sistemas y de los libros publicados sobre dicha materia.

Uno de los principales problemas fue como suministrar índices master para su mantenimiento independiente, si continuamente se estaba cambiando de documentación técnica añadida a los manuales de sistema por parte de los proveedores de los sistemas.

El primer intento para dar solución a este problema fue irónicamente llamado SOFABED (Standar Open Formal Architecture for Browsable Electronic Documents).

El problema de proporcionar índices master vivos era tan fascinante que en 1993, se creó un nuevo grupo CapH (Conventions for the Application HyTime) que aplicaría las sofisticas características del hipertexto en la norma ISO10774, más conocida como standard Hytime, cuyo principal objetivo fuera posibilitar la fusión de índices impresos.

Posteriormente evolucionó hacia otras estructuras (como tesauros), hasta llegar a ser una herramienta considerada en la web para la organización, representación y gestión del conocimiento. La primera versión oficial de estándar ISO/IEC data del año 2000.

En la práctica, la implementación de los topic map se diseñó para arquitectura SGML con notación de HyTime, la DTD del estándar ISO/IEC 13250:2000 se

ha mantenido en HyTime. No obstante, esta notación ha caído en desuso tras la aparición de XML. En este caso, las causas del desplazamiento de la notación de HyTime se deben a la creación de una DTD para crear topic maps en XML, denominada XTM, y al auge que ha obtenido XML. En cualquier caso, actualmente se pueden encontrar ejemplos de ambos lenguajes en la Web.

2.3.2. Definición de TM

Topic Maps es un concepto que se está desarrollando en el entorno web como herramienta que facilita la navegación entre conceptos. Un TM tiene como finalidad estructurar la información mediante la construcción de una red de enlaces semánticos que relacionan diferentes recursos informativos, de tal manera que se pueda realizar la representación de conocimiento.

El estándar Topic Maps está definido para SGML, y un topic map es un documento o un conjunto de documentos SGML o XML, interrelacionados en un espacio multidimensional en el que las localizaciones son topic.



Figura II.5. Diagrama TAO (Topics Asociaciones Ocurrencias) de un Topic Map

Los conceptos clave que forman un topic map son:

- **Topic:** es la representación en el topic map de un concepto abstraído de una realidad cualquiera. La abstracción puede ser individual, es decir puede referirse a sujetos particulares o puede hacer referencia a sujetos generales.
- **Topic type:** representa la relación que se establece entre una clase y sus instancias, que son los topics. Aunque los topics sean considerados como instancias de los topics types, éstos a su vez pueden ser considerados como topics.

Ejemplo definición de Topic type

```
<topic id="tm">
  <instanceOf>
    <topicRef xlink:href="#representación conocimiento"/>
  </instanceOf>
  <baseName>
    ...
  </occurrence>
</topic>
```

Figura II.6. Ejemplo de Topic Type

- **Topic occurrence:** son los links a otros recursos informativos en los que aparecen los ejemplos citados, y que, aunque ajenos al topic map, mejoran la comprensión de un Topic. Las occurrences se pueden agrupar en occurrences role.
- **Occurrence role:** Las occurrences pueden ser de distintos tipos en función del recurso al que se enlace, es decir, pueden ser páginas web, artículos, monografías, comentarios, etc., esto es a lo que se le conoce como occurrences role.

Ejemplo definición de Occurrence

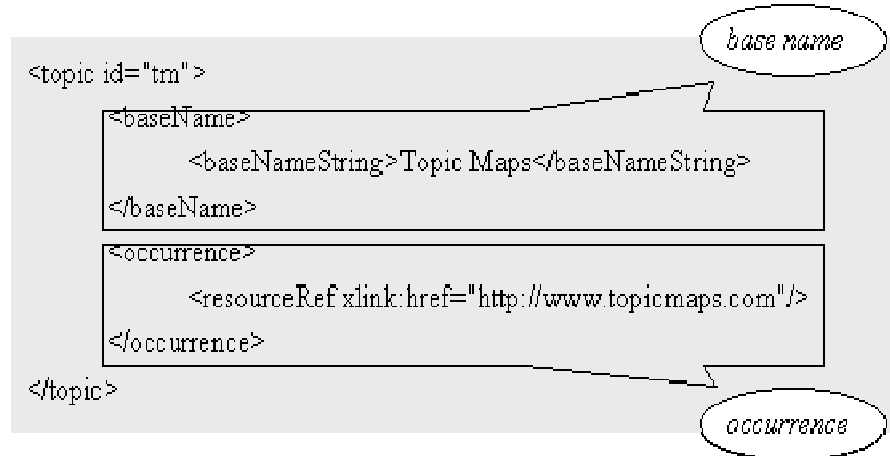


Figura II.7. Ejemplo de Occurrence

Existen dos tipos de ocurrencias: *resourceRef* que es un enlace a un recurso externo de información; y *resourceData* que es algún dato no externo que se facilita (en el siguiente ejemplo, la definición es un resource Data). Cada ocurrencia le corresponde un solo enlace.

```

<topic id="..."><instanceOf>...<baseName></baseName>
  <occurrence>
    <instanceOf><topicRef xlink:href="#definición"/></instanceOf>
    <resourceData>un TM representa el conocimiento</resourceData>
  </occurrence>
  <occurrence>
    <instanceOf><topicRef xlink:href="#webpage"/></instanceOf>
    <resourceRef xlink:href="http://www.topicmaps.com"/>
  </occurrence>
</topic>

```

Figura II.8. Ejemplo de Resource Data

Cuando se diseña un TM es mejor que el número de ocurrencias se limite a unos pocos recursos muy relevantes.

- **Topic association:** es la relación que se establece entre topics. Las topic associations están compuestas por dos o más topics y por las formas

verbales que los une. Las topics association se pueden clasificar según la association type.

- **Association type:** está definida por la forma verbal que une a los topics, es decir, la association type es la que define el verbo que une en cada caso a los topics; es decir una asociación describe una relación entre dos o más tópicos; es un elemento de vínculo.

Por ej.: “*está ubicado en*” y “*se localiza en*” son *associations* que podrían agruparse en un *association type* de “**ubicación**”.

No hay normalización aún respecto de los tipos de asociaciones que pueden usarse, pero generalmente se trata de relaciones del estilo:

- ✓ todo – parte
- ✓ relaciones específicas (“influencia a”, “está cerca de”)
- ✓ relaciones físicas (“está debajo de”, “está cerca de”)
- ✓ relaciones temporales (antes, después)
- ✓ relaciones lógicas (causa – efecto)
- ✓ relaciones jerárquicas

Existen otros elementos que disminuyen la ambigüedad, así por **Association roles** se hace referencia al rol que desempeña determinado topic en una association. En la frase “En Cádiz encontramos pinsapos” se tiene que “Cádiz” desempeña un rol de lugar y el término “pinsapo” designa un tipo de árbol.

Ejemplo de *Association roles*: En una asociación, los topics se denominan miembros y los miembros representan roles, por ejemplo de la frase “Pedro da clase a José en la academia” se puede obtener la siguiente representación ternaria:

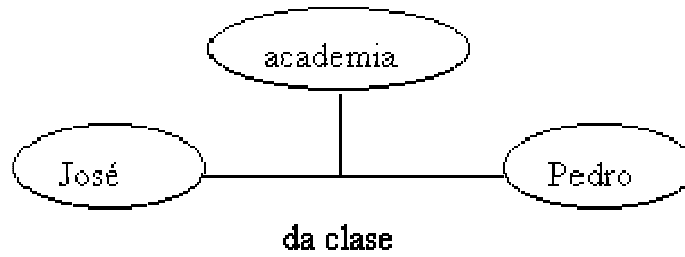


Figura II.9. Ejemplo de Association

Sin añadir más información no se puede saber quién es el profesor y quién el alumno. Esta información se puede expresar mediante roles:

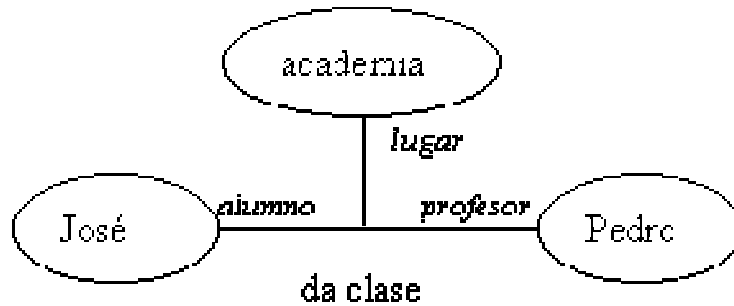


Figura II.10. Ejemplo de Association Role

Como anteriormente, los roles "alumno", "lugar" y "profesor" deberán ser definidos como topics, al igual que "José", "Pedro" y "da clase"

La forma de expresarlo en XML es la siguiente:

```

<association>
  <instanceOf><topicRef xlink:href="#da-clase-a"/></instanceOf>
  <member>
    <roleSpec><topicRef xlink:href="#alumno"/></roleSpec>
    <topicRef xlink:href="#Jose"/>
  </member>
  <member>
    <roleSpec><topicRef xlink:href="#ugar"/></roleSpec>
    <topicRef xlink:href="#academia"/>
  </member>
  <member>
    <roleSpec><topicRef xlink:href="#profesor"/></roleSpec>
    <topicRef xlink:href="#Pedro"/>
  </member>
</association>

```

Figura II.11. Ejemplo de Association en XML

- **Theme/scope:** El scope lo constituyen los límites de validez de las características asignadas a un topic. El límite de la validez de las asignaciones de cada topic se expresa mediante el conjunto de temas (theme) en los que las asignaciones tengan lugar. Un ejemplo podría ser "blanco" en el tema "financiero" o en el tema "pesca".
- **Facet:** Las facets proporcionan un mecanismo para asignar pares de property - value de recursos o fuentes de información. Una facet es simplemente una propiedad, y sus valores son denominados facet values. Las facets son usadas normalmente para suplir la clase de metadatos que pueden ser provistos por atributos SGML o XML. Pueden incluir propiedades como "lenguaje", "seguridad", "aplicabilidad", etc. Las facets pueden ser usadas también para cubrir las clases de propiedades usadas en los sistemas de clasificación facetada. Por último, las facets constituyen un complemento al scope, mientras que el scope puede ser visto como un mecanismo de filtrado que está basado en las propiedades de los topics, las facets proporcionan un filtro basado en las propiedades de las fuentes de información mismas. Las facets son ortogonales al

modelo del topic maps mismo (excepto cuando facets y facets values son vistos como topics).

- **Public subject:** Algunas veces el mismo constructor es representado por más de un topic link. Esto ocurre cuando se mezclan dos topics maps (por ejemplo, "France" y "Francia" pueden ser dos topics distintos que en realidad representan el mismo subject). En estos casos hay que tener alguna forma para establecer la identidad entre topics aparentemente diferentes. El concepto que posibilita esta identidad es el de public subject, y el mecanismo usado es un atributo (el identity attribute) en el elemento topic. Estos atributos se dirigen a un recurso o fuente electrónica que identifica al subject e cuestión de la manera menos ambigua posible. Esta fuente puede ser oficial, un documento validable públicamente (por ejemplo una norma ISO) o puede ser simplemente una definición descriptiva dentro (o fuera) de uno de los topic map considerados.

El concepto más interesante a resaltar en el topic map es el de association type. Un topic map no establece un número concreto y cerrado de tipos posibles de relación, sino que deja abierta la posibilidad de establecer un número potencialmente infinito de tipos de relaciones.

2.3.3. Modelos de TM

2.3.3.1. TMQL (Topic Maps Query Language)

Los Topic Maps se almacenan actualmente en una gran variedad de tecnologías de almacenamiento, como son documentos de texto, documentos XML, bases de datos en XML, bases de datos relacionales y por supuesto bases de datos nativas para Topic Maps. Mientras todos estos soportes tecnológicos ofrecen lenguajes propios como XPATH o SQL para acceder a partes del contenido de los Topic Maps, la necesidad de una recuperación y de un lenguaje de interrogación propia para Topic Maps fue reconocida.

Las aplicaciones pueden usar un procesador **TMQL** para procesar las sentencias del lenguaje de interrogación, que serán evaluadas en el contexto de uno o más Topic Maps

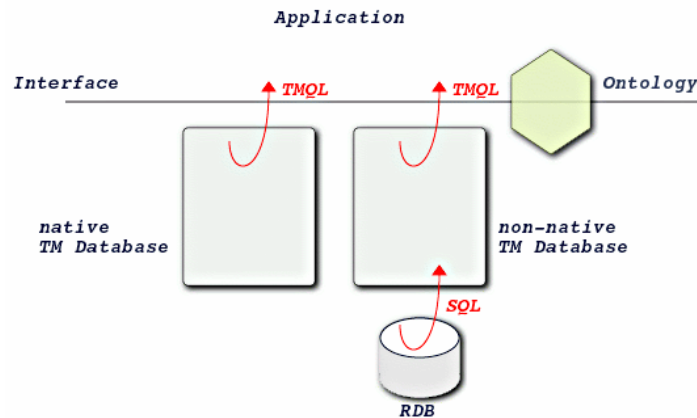


Figura II.12. Ejemplo de TMQL

Topic Maps Query Language (TMQL) definirá un lenguaje para consultar el modelo de datos de mapa conceptual, que permita seleccionar los modelos de mapas conceptuales (como temas y asociaciones) y los datos contenidos en ellos (nombre del tema o valores de ocurrencia, por ejemplo).

2.3.3.2. TMCL (Topic Maps Constraint Language)

Topic Maps Constraint Language (TMCL) define el lenguaje del sistema para los mapas conceptuales que permitirán al autor del sistema limitar los modelos que pueden aparecer en un mapa conceptual y la manera que deben relacionarse entre ellos. Al igual que con XML un lenguaje del sistema para los mapas conceptuales permite la validación y también aplicaciones de edición orientadas al servicio más inteligentes.

2.3.3.3. XTM (XML Topic Maps)

El propósito de un mapa de tópicos es transmitir el conocimiento contenido en recursos a través de una capa sobrepuesta, o mapa, de los propios recursos. Un mapa de tópicos capta los conceptos de los que tratan los recursos, y las

relaciones entre conceptos, en una forma en que es independiente de la implementación.

Los conceptos clave en los mapas de tópicos son tópicos, asociaciones, y ocurrencias.

Un tópico es un recurso situado dentro del ordenador que sustituye a (o "simboliza") algún concepto del mundo real. Ejemplos de dichos conceptos pueden ser la obra *Hamlet*, el autor William Shakespeare, o la relación "autoría".

Los tópicos pueden tener nombres. También pueden tener *ocurrencias*, es decir, recursos de información que se consideran relevantes de algún modo para su concepto. Por último, los *tópicos* pueden participar en relaciones, denominadas *asociaciones*, en las cuales tienen *roles* como miembros.

Así, los tópicos tienen tres tipos de características: nombres, ocurrencias, y roles jugados como miembros de asociaciones. La asignación de dichas características se considera que son válidas para un cierto alcance, o contexto.

Los mapas de tópicos pueden ser combinados en uno solo. La unión puede tener lugar a discreción del usuario o de la aplicación (durante su procesamiento), o puede ser indicada por el autor del mapa de tópicos en el momento de su creación.

2.3.3.4. HyTM

En el verano de 1999 se aceptó como norma y se publicó la ISO/IEC 13250:2000 Topic Maps (International Organization for Standardization, 2000). Está descrito en lenguaje HyTime (Hypermedia/Time-based Structuring Language) el cual, a su vez, es una aplicación del lenguaje SGML (Standard Generalized Markup Language) para procesar documentos. De forma coloquial se conoce como **HyTM**. Esta sintaxis de intercambio es muy flexible lo que permite a los usuarios escribir sus propias DTDs y el uso de

una gran variedad de estructuras de enlace multimedia definidas mediante HyTime (AHMED, 2003). Sin embargo, el número de personas que dominan este lenguaje no es elevado pues a pesar de su flexibilidad, o a causa de ello, resulta complejo.

En la segunda edición de la norma ISO 13250, presentada en 2003, que incorporaba las sintaxis: la primera basada en HyTM y la formulada en la especificación *XTM*, que utilizaba el xml.

2.3.4. Funcionamiento

Recuperación de información con TM en la web

Navegabilidad e inferencia

La navegación tradicional en la web se realiza mediante hiperenlaces que asocian unos recursos informativos a otros, es decir se realiza entre occurrences. Uno de los principales beneficios de los TM reside en la navegabilidad, con TM se puede navegar mediante hiperenlaces entre topics que no tengan un recurso asociado.

No obstante, la navegabilidad entre topics (o occurrences) relacionados no tiene un sentido lógico, dicho de otro modo, por si misma la asociación que se establece entre Cádiz y Pinsapos en la frase "en Cádiz se encuentran Pinsapos" sólo afirma que Cádiz está relacionado con los Pinsapos y los Pinsapos con Cádiz, y que desde cualquiera de los dos elementos se puede acceder ("navegar") hacia el otro.

Otro aspecto interesante de la utilización de TM en la Web es la inferencia. Es decir los mecanismos por los que podemos obtener una información no explícita en el TM.

Las inferencias que se pueden realizar a partir de un conjunto de asociaciones están marcadas por las propiedades que tienen las association type implicadas. Aunque en la literatura se mencionan distintas propiedades como la reflexiva,

conectiva, etc. (Rath, 1999), se habla más de la transitiva y simétrica por ser las más empleadas para inferencia y creación de TM.

Visualización

Los TM representan una de las principales propuestas para la visualización de la web semántica (Le Grand, 2002). Existe una problemática obvia en este tema cuando se tiene en cuenta que un TM puede tener cientos de miles de asociaciones de diferente tipología (association types, roles, occurrences, etc). Actualmente, entre las representaciones propuestas destacan los árboles, los browsers y los gráficos. A la hora de analizar estas representaciones se debe tener en cuenta que muestren tanto información local de los topics que interesan al usuario, como información sobre la localización de esos topics en el conjunto del TM. Por ejemplo:

Árbol con hipervínculos: árbol con hiperenlaces. Uno de los principales inconvenientes de este planteamiento, es que si bien la percepción del TM es más sencilla para el usuario, no pasa lo mismo con los distintos elementos del TM. La tipología de los TM (topics, occurrences, roles, types) hace que el usuario pueda desorientarse ante un árbol de estas características.

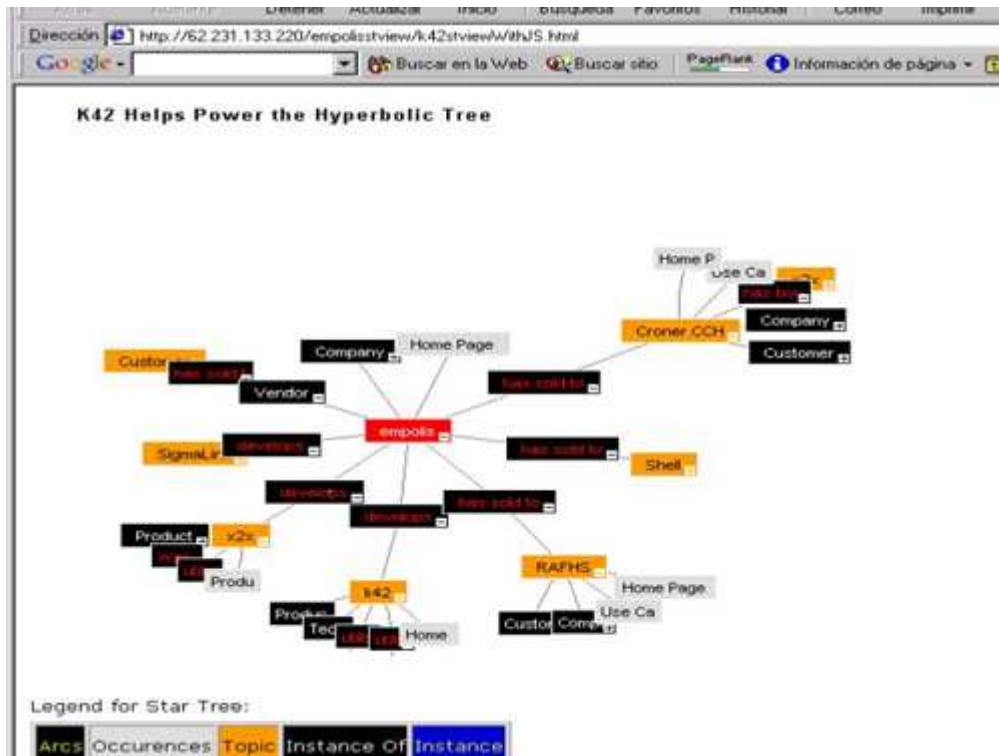


Figura II.13. *Árbol con Hipervínculos o Árbol con hiperenlaces*

Browser: Es decir como un árbol de directorios del Web.

Gráfico: se muestra un nodo principal y los nodos más próximos.

2.3.5. Ventajas y Desventajas

Las **ventajas** que tiene topic map para representar la web semántica se pueden resumir en:

- Los topic maps, pueden dar semántica a elementos que están en el web al organizarlos y describirlos, pero sin modificarlos.
- Perfiles de usuarios: mediante el scope y el theme permite adaptarse a distintas comunidades compartiendo recursos informativos.
- Navegabilidad e inferencia mediante estructuras semánticas. Lo cual mejora no solo la recuperación de información, sino también la gestión del conocimiento y el mantenimiento de los topic maps. En este punto es

también interesante recordar la independencia de los recursos informativos que tiene topic maps, pudiendo navegar por nodos que no tengan ningún recurso asociado.

- Fusión con otras estructuras de conocimiento, permitiendo una gestión descentralizada.
- Buena escalabilidad y compatibilidad para adaptarse al creciente número de recursos de información.

Las posibles **desventajas** residen más en la poca madurez del estándar que en otros aspectos. Así se han revisado distintos aspectos, como:

- La necesidad de mejorar la capacidad de integrar las propiedades de las association types dentro del esquema de definición de topic maps.
- La línea de investigación emergente sobre la forma de asegurar la consistencia de los topic maps.
- Las restricciones necesarias relativas a su validación y corrección.

CAPÍTULO III

REPRESENTACIÓN DEL CONOCIMIENTO BAJO XML

TOPIC MAPS XTM

3.1. Evolución de XTM

Los topic maps se desarrollaron a finales de la década de los 90 como uno de los paradigmas más prometedores para superar los problemas de la organización de grandes espacios de información digital. El web semántico, tal y como fue soñado por Berners-Lee, había abierto el camino para buscar formas de representar la información que fuesen capaces de incorporar no sólo el contenido, sino también la semántica inherente a ellos. En realidad, desde una perspectiva histórica, la búsqueda de la semántica será recordada como el punto de inflexión clave para los procesos de tratamiento de la información. Y dentro de este contexto es donde cabe

situar a los *Topic Maps*. Sin embargo, es preciso aclarar previamente una cuestión que lleva a confusión. Los *Topic Maps* son un estándar *ISO*, concretamente la norma *ISO 13250 Topic Maps*, cuya primera versión fue presentada en el año 2000. Venía derivada de la *ISO/IEC 10744 Hypermedia/Time-based Structuring Language (HyTime)* (1997). Su primera versión establecía el paradigma y los elementos de los *Topic Maps*, usando para ello el lenguaje *HyTM*. Desde el punto de vista de un desarrollador web, *HyTime* utiliza una técnica conocida como arquitectura SGML, recogida en el anexo 3 de la norma *ISO 10744:2* (1997) de alta complejidad y, por lo tanto, de costoso proceso de aprendizaje. Evidentemente, la formulación en SGML no se ajustaba a las necesidades de un contexto en el cual el lenguaje de etiquetado XML se estaba conformando como *lingua franca*. En consecuencia, la comunidad de trabajo organizada alrededor de los *Topic Maps* procedió a proponer una especificación formal para la sintaxis de los *Topic Maps*, que utilizaba XML, y que recibió la denominación de *XTM (Xml Topic Maps)* o lo que es lo mismo, la creación de una DTD para crear *Topic Maps* en XML. Esta especificación formal, en su versión 1.0, se presentó en la conferencia *XML GCA 2000*, en Washington (EUA). Como consecuencia, las herramientas y plataformas de desarrollo que trabajaban con *Topic Maps* para la representación y el acceso a la información adoptaron la especificación *XTM* en lugar de la norma *ISO*. Como ofrecía bastantes ventajas debido a la versatilidad y sencillez de su plantilla se intentaba asegurar que todas las versiones actuales de software comercial que hubiera en el mercado soportaran ambas sintaxis, tanto la *XTM* como la *HyTM*, ya que si se llevaba a cabo una comparativa de las dos quedaba totalmente claro que existían claras e importantes diferencias entre ellas. Esta diferenciación llevó a que se redactara una segunda edición de la norma *ISO 13250*, presentada en 2003, que incorporaba ambas sintaxis: la primera basada en *HyTM* y la formulada en la especificación *XTM*, que utilizaba el XML. Sin embargo, esta segunda versión ha dejado en el aire varios problemas que han llevado al grupo de trabajo *JTC 1 / SC 34 WG 3* de *ISO* a revisar la norma y a proponer una división de la misma en cinco partes completando las lagunas que ofrece. Además, se está trabajando en dos nuevas normas: la *ISO 18048 (Topic Maps Query Language, TMQL)* y la *ISO 19756 (Topic Maps Constraint Language, TMCL)*.

La complejidad cada vez mayor que van a ofrecer los recursos de información digital se refleja en las arquitecturas que deben soportarlos. De hecho, el esfuerzo de la *ISO* por la elaboración de estándares al respecto sobre *Topic Maps* abarca tres capas:

—De modelado: comprende aspectos como el modelo de referencia (Rm), y el de aplicación el estándar (SAM).

—Sintáctica: engloba las DTDs y la documentación necesaria para su construcción en *HyTime*, en *XTM* y unas pautas para conseguir la sintaxis adecuada.

—De restricciones y consultas: incluye *TMQL* y *TMCL*.

3.2. Definición de XTM

Un Topic Map tiene como finalidad normalizar los elementos y la notación utilizada para estructurar la información mediante la construcción de una red de enlaces semánticos que relacione diferentes recursos informativos.

Teóricamente, un Topic Map equivale a un índice creado para una colección de recursos disponibles en formato electrónico, cuyos elementos son:

- Una serie de tópicos o materias que se utilizan para describir el contenido de los recursos y facilitar su recuperación.
- Relaciones que establecen entre los tópicos.
- Una serie de recursos que se indizarán o describirán mediante los tópicos anteriores ello equivale a decir que el recurso trata de o contiene información sobre ese tópico.

Los Topic Maps pueden dar una semántica a elementos que están en el web, al organizarlos y describirlos, pero sin modificarlos.

- Presencia de perfiles de usuarios, permiten adaptarse a distintas comunidades para compartir recursos de información.
- Navegabilidad e inferencia mediante estructuras semánticas: Mejora no sólo la recuperación de información, sino también la gestión del conocimiento y el mantenimiento de los Topic Maps. En este punto, es también interesante recordar la independencia de los recursos de

información que tiene un Topic Map, al poder navegar por nodos que no tengan ningún recurso asociado

- Fusión con otras estructuras de conocimiento, que permite una gestión descentralizada.
- Buena escalabilidad y compatibilidad para adaptarse al creciente número de recursos de información.

Objetivos

Los objetivos de diseño de XTM son:

1. XTM deberá ser directamente utilizable en Internet.
2. XTM deberá soportar una amplia variedad de aplicaciones.
3. XTM deberá ser compatible con XML, XLink, y la ISO 13250.
4. Deberá ser fácil escribir programas que procesen documentos XTM.
5. El número de características opcionales en XTM debe mantenerse en el mínimo absoluto, idealmente cero.
6. Los documentos XTM deberán ser legibles para las personas y razonablemente claros.
7. El diseño XTM deberá ser preparado rápidamente.
8. El diseño de XTM deberá ser formal y conciso.
9. Los documentos XTM deberán ser fáciles de crear.
10. La brevedad del mercado XTM tiene una importancia mínima.

Documento XTM

Limitación de denominación de tópico

El límite, impuesto por el paradigma "mapa de tópicos", de que los tópicos que tienen el mismo nombre base en el mismo alcance se refieren implícitamente al mismo concepto y, por tanto, deben unirse.

Tópico

Un recurso que actúa como sustituto (proxy) para algún concepto; el sistema de representación del mapa de tópicos de ese concepto. La relación entre un tópico y su concepto está definida como una relación de simbolización. La simbolización de un concepto permite asignar características de tópico al tópico que lo simboliza.

Un elemento <topic>.

Nombre de tópico

1. Un nombre base característico de un tópico (incluidas las variantes del nombre base).
2. (*Informalmente*) La cadena de caracteres especificada como nombre de un tópico usando un elemento <baseNameString>.

Característica de tópico

Cualquier cosa que pueda afirmarse sobre un tópico en el paradigma mapa de tópicos se conoce como característica de ese tópico. Pueden ser características uno de los siguientes:

- Un nombre de tópico
- Una ocurrencia de tópico
- Un rol jugado por un tópico como miembro de una asociación

La asignación de tales características se considera válida dentro de un cierto alcance, o contexto.

Los nombres de tópico, las ocurrencias, y los roles jugados en las asociaciones se conocen colectivamente como características.

Mapa de tópicos

Un mapa de tópicos es una colección de tópicos, asociaciones, y alcances (llamados colectivamente nodos del mapa de tópicos) que pueden existir en una de dos formas:

1. Un formato de intercambio serializado (por ejemplo un documento mapa de tópicos expresado en sintaxis XTM), o
2. Alguna forma de aplicación interna.
3. El elemento documento mapa de tópicos (<topicMap>) expresado usando la sintaxis XTM.

El propósito de un mapa de tópicos es transmitir el conocimiento contenido en recursos a través de una capa sobrepuesta, o mapa, de los propios recursos. Un mapa de tópicos capta los conceptos de los que tratan los recursos, y las relaciones entre conceptos, en una forma en que es independiente de la implementación.

Ocurrencia

Una ocurrencia es cualquier información considerada relevante para un concepto dado. Las ocurrencias constituyen uno de las tres clases de característica que pueden ser asignadas a un tópico y están, por tanto, gobernadas por el alcance. Cada ocurrencia individual es una instancia de una sola clase de ocurrencia (también conocida como tipo ocurrencia) que puede o no ser indicada explícitamente. El tipo ocurrencia por defecto está definido por el concepto publicado "ocurrencia".

Un elemento hijo (<occurrence>) de un elemento <topic>.

Una ocurrencia de tópico.

Ocurrencia de tópico

Un recurso que contiene información considerada relevante para un concepto dado. Para ser expresado en un mapa de tópicos XTM, dicho recurso puede ser:

- Direccionado vía una URI⁴ usando un elemento <resourceRef>, o
- Situado en el documento XTM como un elemento <resourceData>.

⁴ *Uniform Resource Identifier*

El último (datos de recurso) proporciona una forma útil de expresar una pieza corta de información acerca de un concepto (por ejemplo, la fecha de composición de una obra).

Asociación

Una asociación es una relación entre uno o más tópicos, cada uno de los cuales juega un rol como miembro de esa asociación. Los roles que un tópico juega en las asociaciones están entre las características que pueden serle asignadas y, por tanto, son gobernadas por el alcance. Cada asociación individual es una instancia de una sola clase de asociación (también conocida como tipo asociación) que puede o no ser indicada explícitamente. El tipo asociación por defecto está definido por el concepto publicado "asociación".

No hay direccionalidad inherente en una asociación. (Las asociaciones describen relaciones: Si A está relacionado con B, entonces B debe también estar relacionado con A. Está más relacionado con cuál es la clase de relación y que papeles son jugados por sus miembros. La cuestión de cómo marcar una relación es un problema de denominación, no de dirección.)

Miembro

1. Un elemento hijo (<member>) de un elemento <association>.
2. Un conjunto de tópicos que juegan un determinado rol en una asociación.

Clase-Instancia

Clase-instancia es una clase de asociación que expresa las relaciones clase-instancia entre tópicos que juegan los roles de clase e instancia respectivamente. Los conceptos "clase-instancia", "clase", e "instancia" están todos definidos por indicadores de concepto publicados (PSI⁵s)

Superclase-Subclase

Superclase-subclase es una clase de asociación que expresa las relaciones superclase-subclase entre tópicos que juegan los roles de superclase y subclase respectivamente. Los conceptos "superclase-subclase",

⁵ *public subject indicator*

"superclase", y "subclase" están todos definidos por indicadores de concepto publicados (PSIs).

3.3. Funcionamiento de XTM

Introducción a la sintaxis XTM

La sintaxis para serializar e intercambiar documentos mapa de tópicos conformes a esta especificación está descrita en la definición de tipo de documento XML (DTD XML) aclarada más adelante en Normativas Actuales. Esta sección proporciona documentación para todos los elementos definidos en esa DTD.

A continuación se da una lista completa de los tipos de elementos XTM en el orden en el que se documentan:

- <topicRef>: referencia a un elemento tópico
- <subjectIndicatorRef>: referencia a un indicador de concepto
- <scope>: referencia al tópico(s) que determina el alcance
- <instanceOf>: Apunta a un tópico que representa una clase
- <topicMap>: elemento documento mapa de tópicos
- <topic>: elemento tópico
- <subjectIdentity>: concepto simbolizado por un tópico
- <baseName>: nombre base de un tópico
- <baseNameString>: recipiente de la cadena de caracteres del nombre base
- <variant>: formas alternativas del nombre base
- <variantName>: recipiente de la variante del nombre
- <parameters>: contexto de procesamiento para las variantes
- <association>: asociación de tópicos

- <member>: miembro en la asociación de tópicos
- <roleSpec>: Apunta a un tópico que actúa como rol en la asociación
- <occurrence>: recursos considerados como ocurrencia
- <resourceRef>: referencia a un recurso
- <resourceData>: recipiente de recursos de datos
- <mergeMap>: Unión con otro mapa de tópicos

3.4. Ventajas y Desventajas de XTM

Topic Map and XSLT for the WEB

Ventajas

- Plataforma y software independiente.
- XSLT mecanismos de encapsulación (XML Web Applications Template Library - XWATL)
- Facilidad, simplicidad y belleza
- Escribir una vez traducir en cualquier lugar
- Facilitar la estructuración de la información
- De acuerdo a como se ve y se siente totalmente los sitios web

Entre las posibles **desventajas**, residen más en la poca madurez del estándar que en otros aspectos. Así se han revisado distintos aspectos como:

- La necesidad de mejorar la capacidad de integrar las propiedades de las association types dentro del esquema de definición de *Topic Maps*.
- La línea de investigación emergente sobre la forma de asegurar la consistencia de los *Topic Maps*.
- Las restricciones necesarias relativas a su validación y corrección.

3.5. Normativas actuales

XTM 1.0 Definición del Tipo de Documento (Normativa)

Nota: En la versión online de esta DTD, el nombre de cada tipo de elemento de cada declaración de elemento enlaza con documentación en esta especificación. Los nombres de cada tipo de elemento en los modelos de contenido enlazan con sus respectivas declaraciones en la DTD.

```

<!-- archivo: xtml.dtd -->
<!-- XML Mapas de tópicos (XTM) DTD, Versión 1.0
Este documento es XTM, una sintaxis de intercambio en
XML para la norma
ISO 13250 Topic Maps.
XML Topic Map (XTM)
Copyright 2000-2001 TopicMaps.Org, Todos los derechos
reservados.

Se garantiza aquí, a perpetuidad, el permiso para usar,
copiar, modificar y distribuir la DTD XTM y sus
materiales adjuntos para cualquier propósito y sin
tasas, a condición de que el anterior copyright y este
párrafo aparezcan en todas las copias.
El titular del copyright no se hace responsable de la
idoneidad del DTD para ningún uso que de ella se haga.
Se distribuye "tal cual" sin garantía expresa o
implícita.

Editores: Steve Pepper
<pepper@ontopia.net>
Graham Moore
<gdm@empolis.co.uk>
Autores: Murray Altheim
<altheim@eng.sun.com>
Michel Biezunski
<mb@infoloom.com>
Sam Hunting
<shunting@etopicality.com>
Steven R. Newcomb
<srn@coolheads.com>

```

```

Estatus:      Lanzamiento
Versión:      v1.0.1
Revisión:     $Id:   xtml.dtd,v   1.2   2001/02/08
16:03:12 pepper Exp $

Traducción al español: Ma Jesús Colmenero Ruiz
<mcolmene@bib.uc3m.es>

Identificador público: (esta versión en español no
lo posee aún)

Revisiones:
#2001-01-21: baseName eliminado de occurrence
#2001-02-02: variantName hecho opcional en variant
#2001-02-02: cambiado ID a #IMPLIED en association
#2001-02-02: cambiado ID a #IMPLIED en resourceData
#2001-02-02: cambiado PLUS a REP en member
-->

<!-- Use este URI para identificar el espacio de nombre
XTM por defecto:
      "(esta DTD en español no posee URI)"
      Usado para identificar el espacio de nombre
XLink:
      "http://www.w3.org/1999/xlink"
-->

<!-- topicMap: elemento documento Mapa de tópicos
..... -->
<!ELEMENT topicMap ( topic | association | mergeMap )*>
<!ATTLIST topicMap
      id          ID          #IMPLIED
      xmlns       CDATA       #FIXED 'http://(esta DTD
en español no posee URI)'
      xmlns:xlink CDATA       #FIXED
'http://www.w3.org/1999/xlink'
      xml:base    CDATA       #IMPLIED
>

```

```

<!--          topic:          elemento
Topic.....
-->
<!ELEMENT topic ( instanceOf*, subjectIdentity?, (
baseName | occurrence )* )>
<!ATTLIST topic
      id          ID          #REQUIRED
>

<!-- instanceOf: Apunta a un Topic representando una
clase..... -->
<!ELEMENT instanceOf ( topicRef | subjectIndicatorRef )
>
<!ATTLIST instanceOf
      id          ID          #IMPLIED
>

<!-- subjectIdentity: Concepto representado por el Topic
..... -->
<!ELEMENT subjectIdentity ( resourceRef?, ( topicRef |
subjectIndicatorRef )* )>
<!ATTLIST subjectIdentity
      id          ID          #IMPLIED
>

<!--      topicRef:      Referencia      a      un      elemento
Topic..... -->
<!ELEMENT topicRef EMPTY >
<!ATTLIST topicRef
      id          ID          #IMPLIED
      xlink:type  NMTOKEN  #FIXED 'simple'
      xlink:href  CDATA    #REQUIRED
>

<!-- subjectIndicatorRef: Referencia a un indicador de
Subject..... -->
<!ELEMENT subjectIndicatorRef EMPTY >
<!ATTLIST subjectIndicatorRef
      id          ID          #IMPLIED

```

```

    xlink:type      NMTOKEN   #FIXED 'simple'
    xlink:href      CDATA     #REQUIRED
  >

  <!-- baseName: Nombre Base o preferente de un Topic
  ..... -->
  <!ELEMENT baseName ( scope?, baseNameString, variant* )
  >
  <!ATTLIST baseName
    id              ID        #IMPLIED
  >

  <!-- baseNameString: Depósito de la cadena de caracteres
  del nombre base..... -->
  <!ELEMENT baseNameString ( #PCDATA ) >
  <!ATTLIST baseNameString
    id              ID        #IMPLIED
  >

  <!-- variant: Formas alternativas del nombre
  base..... -->
  <!ELEMENT variant ( parameters, variantName?, variant* )
  >
  <!ATTLIST variant
    id              ID        #IMPLIED
  >

  <!-- variantName: Depósito de la variante del
  nombre..... -->
  <!ELEMENT variantName ( resourceRef | resourceData ) >
  <!ATTLIST variantName
    id              ID        #IMPLIED
  >

  <!-- parameters: Contexto de proceso para el elemento
  Variant ..... -->
  <!ELEMENT parameters ( topicRef | subjectIndicatorRef
  )+ >
  <!ATTLIST parameters

```

```

    id          ID          #IMPLIED
  >

  <!-- occurrence: Recursos considerados como una
  Occurrence ..... -->
  <!ELEMENT occurrence ( instanceOf?, scope?, (
  resourceRef | resourceData ) )>
  <!ATTLIST occurrence
    id          ID          #IMPLIED
  >

  <!-- resourceRef: Referencia a un recurso
  ..... -->
  <!ELEMENT resourceRef EMPTY >
  <!ATTLIST resourceRef
    id          ID          #IMPLIED
    xlink:type  NMTOKEN  #FIXED 'simple'
    xlink:href  CDATA    #REQUIRED
  >

  <!-- resourceData: Depósito de recurso de datos
  ..... -->
  <!ELEMENT resourceData ( #PCDATA )>
  <!ATTLIST resourceData
    id          ID          #IMPLIED
  >

  <!-- association: Relación entre Topics
  ..... -->
  <!ELEMENT association ( instanceOf?, scope?, member+ )>
  <!ATTLIST association
    id          ID          #IMPLIED
  >

  <!-- member: Indicación de pertenencia una relación
  entre Topics ..... -->
  <!ELEMENT member (roleSpec?,(
  topicRef|resourceRef|subjectIndicatorRef )* )>
  <!ATTLIST member

```



```

    id          ID          #IMPLIED
  >

  <!-- roleSpec: Apunta a un Topic que funciona
describiendo el rol
    realizado en la relación (Association
Role)..... -->
  <!ELEMENT roleSpec ( topicRef | subjectIndicatorRef ) >
  <!ATTLIST roleSpec
    id          ID          #IMPLIED
  >

  <!-- scope: Referencia a un Topic que constituye el
contexto de validez ..... -->
  <!ELEMENT scope ( topicRef | resourceRef |
subjectIndicatorRef )+ >
  <!ATTLIST scope
    id          ID          #IMPLIED
  >

  <!-- mergeMap: Unión con otro Mapa de materias
..... -->
  <!ELEMENT mergeMap ( topicRef | resourceRef |
subjectIndicatorRef )* >
  <!ATTLIST mergeMap
    id          ID          #IMPLIED
    xlink:type  NMTOKEN  #FIXED 'simple'
    xlink:href  CDATA    #REQUIRED
  >

  <!--Fin de la DTD XML Topic Map (XTM) 1.0 -->

```

Figura III.1.DTD del XTM

3.6. Software

El software utilizado para el procesamiento del documento XTM, es la herramienta de desarrollo Visual Studio .Net 2008 utilizando el lenguaje C#, los datos abstraídos de la base de datos en SQL Server 2005.

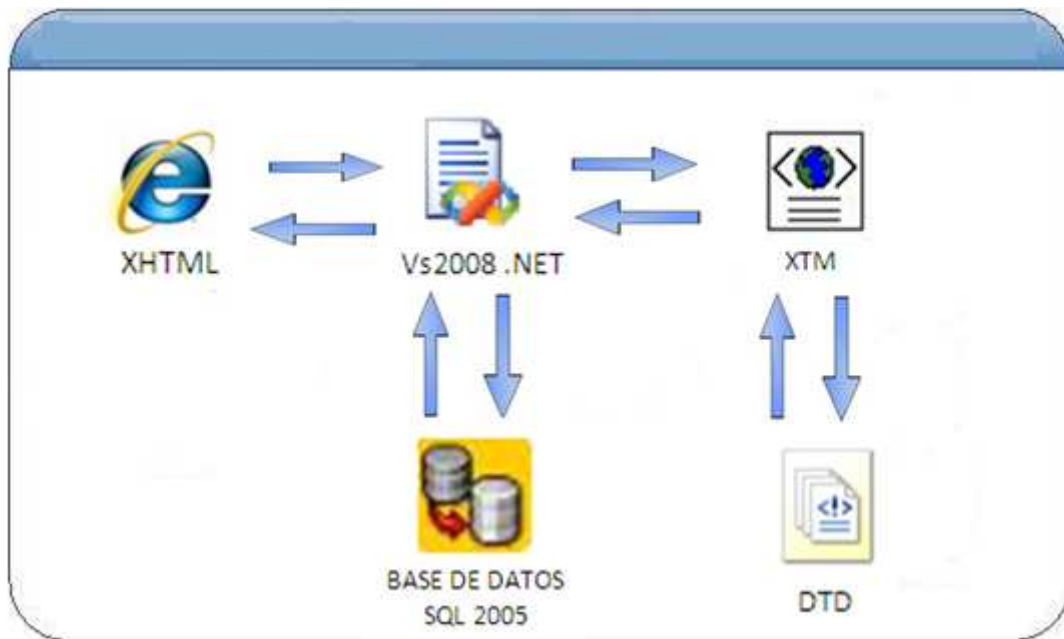


Figura III.2. Funcionamiento de XTM

CAPÍTULO IV

DESARROLLO E IMPLEMENTACIÓN DEL PORTAL SEMÁNTICO

Para la realización del presente proyecto se han seguido las directrices marcadas por la metodología XP (*Extreme Programming*). Es necesaria una metodología ágil como X.P que se adapta a las necesidades del cliente y dónde la aplicación se va reevaluando en periodos cortos de tiempo., además la comunicación es un punto importante y debe realizarse entre los programadores, los jefes de proyecto y los clientes. Esta metodología consta de cuatro fases: fase de análisis, fase de diseño, fase de codificación y fase de pruebas.

4.1. FASE DE ANÁLISIS

ANTECEDENTES

La finalidad del presente documento es describir el área de trabajo, de una forma detallada, ayudar a establecer la viabilidad del esfuerzo del desarrollo, describiendo el entorno sobre el cual se va a realizar la aplicación teniendo en consideración los requerimientos de los usuarios de software.

El plan se concentra en el desarrollo de un portal semántico basado en el paradigma XML Topic Maps (XTM) de La institución Educativa Liceo Stephen Hawking, especificando cómo y cuándo se lo realizará, y una estimación de los costos de operación y mantenimiento.

4.1.1. EXPLORACION DEL ÁMBITO Y DEL PROBLEMA

4.1.1.1. ANTECEDENTES HISTÓRICOS

Es LICEO porque pretendemos configurarnos como un espacio educativo distinto a la escuela tradicional, un lugar en donde los niños y niñas aprendan jugando, sean protagonistas y desarrollen al máximo sus potencialidades; y lleva el nombre de STEPHEN HAWKING, en homenaje al más grande científico.

El Liceo Stephen Hawking, es una nueva institución educativa, de nivel pre básico y básico que está inició sus actividades preparatorias al año lectivo 2001-2002, con pre kinder, primero y segundo básico, es decir con niñas y niños de 4, 5 y 6 años de edad. Se sustenta en un nuevo modelo educativo que busca la formación de niñas y niños con seguridad en sí mismos, reflexivos, críticos y responsables, con alto desarrollo de la inteligencia y comprometidos con la equidad y transformación social.

Las actividades formativas son personalizadas y se hacen en talleres interactivos en aulas con máximo 20 estudiantes. En las tardes se ofrecen talleres aptitudinales de: artes plásticas, creatividad e ingenio, baile y danza, música, internet y ciencias, lectura y deportes. Tareas Dirigidas.

En la actualidad existe la educación hasta séptimo de básica.

Esta innovadora propuesta educativa para desarrollo del pensamiento, se divulgó en un acto de afecto educativo el 1 julio 2001 en el Teatrino de la Casa de la Cultura núcleo de Chimborazo.

La Visión del Liceo Stephen Hawking es:

El liceo ofrece una formación personalizada. La metodología es de talleres y proyectos. Se apoya en talleres aptitudinales de: Ciencias y creatividad, Internet y computación, artes plásticas, deportes, música y danza y baile.

Cuenta los servicios educativos de: médico, odontólogo, nutricionista, psicopedagogo y psicólogo.

La Misión del Liceo Stephen Hawking es:

Nace con la misión de formar niñas y niños seguros, expresivos, autónomos, creativos y reflexivos, con alto desarrollo intelectual, y comprometidos con la equidad y transformación social. El proyecto educativo del liceo se sustenta en 4 ejes:

- Desarrollo del pensamiento.
- Participación infantil
- Desarrollo de la expresión corporal, oral, musical, lúdica y creativa.
- El buen trato, como ambiente de acogida y afecto humano.

A continuación el organigrama

En esta distribución gráfica señalamos la organización estructural administrativa de la Unidad Educativa.



4.1.1.2. OBJETIVOS DEL LICEO STEPHEN HAWKING.

Objetivos de información

- El habito del pensamiento creativo e independiente
- Los talentos artísticos creativos y de liderazgo
- Una imagen de sí mismo positivamente valiosa
- Conocimiento y apreciación de las cosas las personas y los hechos de su ambiente
- Las habilidades escolares básicas de lectura, escritura y cálculo

Objetivos de la Educación Básica

- La educación básica ecuatoriana se compromete a ofrecer las condiciones más apropiadas para los jóvenes, al finalizar este nivel de educación, logre el siguiente perfil:
 - Conciencia clara y profunda del ser ecuatoriano, en el marco del reconocimiento de la diversidad cultural, étnica geográfica y de género del país.
 - Conscientes de sus derechos y deberes de relación a si mismo a la familia, a la comunidad y a la nación.
 - Alto desarrollo de su inteligencia a nivel del pensamiento creativo práctico y teórico.
 - Con capacidad de aprender, con personalidad autónoma y solidaria con su entorno social y natural, con ideas positivas de si mismos.
 - Con actitudes positivas frente al trabajo y al uso del tiempo libre.

4.1.1.3. FUNCIONAMIENTO

El Instituto Liceo Stephen Hawking es el encargado de mostrar su información transparente a usuarios que deseen conocer más sobre la Institución de manera semántica, para gestionar este trabajo se necesitaría automatizar las siguientes tareas:

- Gestión de Información General

- Gestión de Información Semántica

SUBSISTEMA DE GESTIÓN INFORMACIÓN GENERAL

- Dar a conocer misión, visión, historia, objetivos, y generalidades relevantes a la Institución.
- Llevar los registros de información de las autoridades de la institución.
- Mostrar los docentes contratados de la Institución

INTERFACES DE USUARIO

La interfaz de usuario maneja múltiples páginas web, el manejo del portal semántico se lo realizará a través del teclado o del ratón, al estilo de navegación web.

INTERFAZ DE COMUNICACIÓN

La conexión a la red se establece por medio de la intranet de la institución y mediante la utilización de navegadores (Browsers) para Internet.

PROCESO INVOLUCRADO

La obtención de datos proviene de un documento, en donde se encontrará la información requerida y esta es:

- Documentos Microsoft Excel.
- Documentos Microsoft Office.

La información mencionada anteriormente tiene datos relativos a la institución, sobre los cuales se trabajará de acuerdo a los índices planteados y propuestos.

En los documentos Microsoft Excel encontramos los siguientes datos:

- Nombres de los estudiantes de cada paralelo de cada grado

En los documentos Microsoft Office encontramos los siguientes datos:

- Historia de la institución
- Misión
- Visión
- Objetivos

- Información sobre matriculas e inscripciones (requisitos para cada alumno)
- Periodos académicos
- Dirección, Teléfono, email de la institución
- Personal de la institución con sus cargos y datos personales
- Cada docente y que materia dicta con sus datos personales y a que grados imparte que materias (Horario docente)
- Número de grados y paralelos
- Número de estudiantes de cada grado por paralelo
- Número de alumnos total de la institución
- Número de docentes total
- Horario de cada grado

SUBSISTEMA DE GESTIÓN INFORMACIÓN SEMÁNTICA

- Mostrar el pensum que posee la Institución
- Mostrar los Estudiantes que están matriculados en un grado
- Publicar el horario de un grado
- Dar a conocer el curriculum de cada docente contratado
- Mostrar las materias que dicta cada docente

INTERFACES DE USUARIO

La interfaz de usuario utiliza páginas web, y el manejo del sistema se lo realizará a través del teclado o del ratón, al estilo de navegación web.

INTERFAZ DE COMUNICACIÓN

La conexión a la red se establece por medio de la intranet de la institución y mediante la utilización de navegadores (Browsers) para Internet.

PROCESO INVOLUCRADO

Los datos a los cuales haremos referencia para la obtención de las estadísticas residen o provienen de nuestra base de datos que a continuación daremos a

conocer:

- Liceo

La base mencionada anteriormente tiene definidas tablas y sus respectivos campos sobre los cuales realizaremos las respectivas consultas SQL en relación a los índices planteados y propuestos.

La Base de Datos **Liceo** contiene las siguientes tablas:

- Autoridades
- Docentes
- Estudiante
- Grados
- Horario
- Materias

4.1.1.4. ANTECEDENTES TECNOLÓGICOS.

Para la implantación del Portal Semántico del Liceo Stephen Hawking se requiere de una buena infraestructura de soporte que garantice el acceso a la aplicación desde donde se requiera. El Liceo Stephen Hawking no cuenta actualmente con una infraestructura, para lo cual se creará una Base de Datos “Liceo”.

Entre los principales recursos tecnológicos que posee Liceo Stephen Hawking se encuentran:

Red Institucional

- La Institución posee equipos de cómputo.
- La red utiliza principalmente un cableado de cobre.
- Se cuenta con equipos de conmutación que aceptan medios de cobre.

Acceso Satelital a Internet

- WLAN, para acceso a Internet.
- Acceso a Internet desde casi todos los nodos conectados a la red.

Infraestructura de Servidores

- Un servidor Web que utiliza la plataforma WINDOWS SERVER 2003 con IIS y un gestor de base de datos SQL SERVER 2005.

Usuarios

- Existen equipos con características heterogéneas en cuanto a hardware, a través de los cuales se brinda el acceso a Intranet o que se utilizan como estaciones de trabajo aisladas. Todos tienen sistemas operativos compatibles con Win32.

Esta infraestructura tecnológica puede resumirse en el siguiente gráfico:

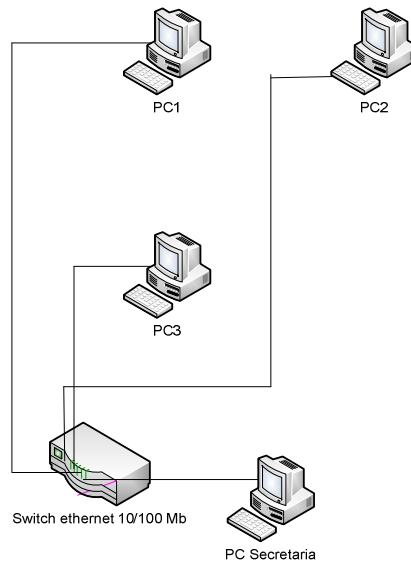


Figura IV.1. Infraestructura tecnología del Liceo Stephen Hawking

4.1.1.5. DEFINICIÓN DE REQUERIMIENTOS DE LOS USUARIOS

TÉCNICAS:

La técnica que se han utilizado para investigar la funcionalidad de Liceo Stephen Hawking son:

- Entrevistas
- Planificación incremental (release)
- Testing
- Diseño simple
- Integración continua

ENTREVISTAS

Cuando se realizó la entrevista a la Directora Lic. María Elena Cisneros Sánchez, y a la Secretaria Mayra Torres, se pudo determinar algunos aspectos importantes que se definió en el punto de funcionamiento.

RELEASES (Historial de Usuarios)

SISTEMA ACTUAL

En el Liceo Stephen Hawking no existe un Portal Semántico, se llevan la información a través de archivos guardados en Microsoft Excel y Microsoft Office.

Historia de Usuario	
Número: 1	Usuario: Secretaria, Directora
Nombre historia: Ingreso de Estudiante.	
Prioridad en negocio: Alta	
Responsable: Secretaria, Directora	
Descripción: La interacción inicia al momento de inscribir o matricular estudiantes, sus datos se guardan en un archivo de Microsoft Excel.	
Observaciones: Para el ingreso de los estudiantes se guardan datos relevantes al mismo.	

Tabla IV.1. Historia de usuario1 del sistema actual

Historia de Usuario	
Número: 2	Usuario: Secretaria, Directora
Nombre historia: Ingreso de Docentes	
Prioridad en negocio: Alta	

Responsable: Secretaria, Directora
Descripción: La interacción inicia contratando docentes, y a continuación se guardan sus datos personales y curriculum en archivos de Microsoft Office
Observaciones: Para el ingreso de los docentes se guardan datos relevantes al mismo.

Tabla IV.2. Historia de usuario2 del sistema actual

Historia de Usuario	
Número: 3	Usuario: Secretaria, Directora
Nombre historia: Ingreso de grado	
Prioridad en negocio: Alta	
Responsable: Secretaria, Directora	
Descripción: La interacción inicia ingresando el grado a continuación del actual	
Observaciones: Para ingresar un grado se verifica que el grado anterior haya tenido éxito	

Tabla IV.3. Historia de usuario3 del sistema actual

Historia de Usuario	
Número: 4	Usuario: Secretaria, Directora
Nombre historia: Ingreso de materias por grado	
Prioridad en negocio: Alta	
Responsable: Secretaria, Directora	
Descripción:	

La interacción inicia ingresando la materia correspondiente a cada grado
Observaciones: Para ingresar una materia se verifica que el grado exista y que un docente pueda dictar esa materia

Tabla IV.4. Historia de usuario4 del sistema actual

Historia de Usuario	
Número: 5	Usuario: Secretaria, Directora
Nombre historia: Ingreso de información de la Institución	
Prioridad en negocio: Alta	
Responsable: Secretaria, Directora	
Descripción: La interacción inicia ingresando información importante de la Institución, misma que es guardada en un archivo de Microsoft Office	
Observaciones: Para ingresar datos importantes de la Institución se toma en cuenta su misión, visión, historia y se plantean objetivos cada determinado tiempo.	

Tabla IV.5. Historia de usuario5 del sistema actual

SISTEMA PROPUESTO

MÓDULO DE INTEGRACIÓN DE BASE DE DATOS

Historia de Usuario	
Número: 1	Usuario: Secretaria, Directora
Nombre historia: Base de Datos	

Prioridad en negocio: Alta
Programador responsable: Pamela Fuertes
Descripción: Es necesario crear la Base de Datos “Liceo”; con la finalidad de automatizar los datos existentes en la Institución y que la información presentada se genere siempre con información actualizada.
Observaciones: Es importante la creación de la Base de Datos en la Institución para facilitar el manejo de información, y para la integración de futuros sistemas

Tabla IV.6. Historia de usuario1 del sistema propuesto

MÓDULO DE BÚSQUEDA SEMÁNTICA DE INFORMACIÓN

Historia de Usuario	
Número:2	Usuario: Cualquier usuario
Nombre historia: Búsqueda Semántica	
Prioridad en negocio: Alta	
Programador responsable: Pamela Fuertes.	
Descripción: Una vez que el usuario ingresa en el Portal Semántico, internamente el portal hace su búsqueda de acuerdo al tema escogido en el documento XML Topic Maps (XTM), y devuelve el o los temas a los que está directamente relacionado.	
Observaciones: Al momento de realizar la búsqueda semántica en el documento XTM, ayuda al usuario en la navegabilidad y es eficiente al mostrar información que proveerá la base de datos.	

Tabla IV.7. Historia de usuario2 del sistema propuesto

MÓDULO DE FUNCIONALIDAD DEL USUARIO

Historia de Usuario	
Número: 3	Usuario: Cualquier usuario
Nombre historia: Información General de la Institución	
Prioridad en negocio: Alta	
Programador responsable: Pamela Fuertes	
Descripción: El usuario ingresa al Portal Semántico y accede a la información general de la institución como es la misión, visión, objetivos, historia, generalidades, autoridades, lista de docentes contratados, y pensum académico.	
Observaciones: Los usuarios que pueden acceder a esta información es cualquier persona interesada en conocer más sobre la Institución. El usuario notará que la navegabilidad en el Portal Semántico es sencilla y eficiente.	

Tabla IV.8. Historia de usuario3 del sistema propuesto

4.1.1.6. DEFINICIÓN DEL PROBLEMA.

IDENTIFICACIÓN DEL ESCENARIO PROBLEMA

El Liceo Stephen Hawking demanda la presencia en la web en el ámbito académico, para representar su información y conocimiento tanto de la institución como de todos los datos importantes de la Institución, implementando un Portal Semántico para alcanzar este objetivo.

Los indicadores serán clasificados por:

- Gestión información general
- Gestión información semántica

ESCENARIOS PROBLEMAS

En la actualidad no se cuenta con ningún tipo de portal que respalde la labor de emitir a los usuarios la posibilidad de notificarse sobre información relevante de la Institución.

A partir de lo descrito en todo el documento presentamos un análisis de los problemas a los cuales se busca dar solución.

Ausencia de un sistemas centralizado.

La falta de una base de datos centralizada para almacenar información proveniente de un sistema de archivos basado en Microsoft Excel y Office, con lo cual no se cuenta con información organizada para la presentación de un portal a la web.

4.1.1.7 ANÁLISIS DE ALTERNATIVAS DE SOLUCIÓN

SOLUCIÓN A NIVEL DE SOFTWARE

OBJETIVOS:

GENERAL

- Utilizar herramientas de desarrollo actual y completas que nos permitan la realización e implementación de un Portal Semántico para representar información y conocimiento para el Liceo Stephen Hawking.

ESPECÍFICOS

- Combinar la eficacia de un portal semántico diseñado para el Liceo Stephen Hawking, con herramientas que son Microsoft Visual Studio.NET 2008, bloque de notas y SQL Server 2005.
- Gestionar en el ámbito de la representación de la información y el conocimiento en beneficio de la institución y su comunidad.

MÓDULO DE INTEGRACIÓN DE BASE DE DATOS

La institución proveerá de la información necesaria para modelar el conocimiento y de esta manera almacenarlos en la base de datos “Liceo”, para que luego puedan ser mostrados en el portal semántico.

MÓDULO DE BÚSQUEDA SEMÁNTICA DE INFORMACIÓN

Los datos se obtendrán a partir de la información que proporcione la Institución con el fin de organizarlos semánticamente en un documento XML Topic Maps (XTM), para mejorar la navegabilidad y eficiencia al momento de ingresar en el Portal de la Institución.

Implementación e Implantación de la Solución:

La implementación del Portal Semántico debe considerar:

- La representación de la información y el conocimiento utilizando el paradigma XTM.
- Las interfaces necesarias para la publicación de la información y el conocimiento generados a través de consultas preestablecidas y diseñadas.
- Además el portal debe considerar parámetros de calidad como navegabilidad, usabilidad, disponibilidad, otras.

MODULO DE FUNCIONALIDAD DEL USUARIO

- Se permite al usuario que visite el Portal Semántico de la Institución sin restricción alguna (no es necesario registrarse, la información es transparente), con la facilidad de no perder el sentido de búsqueda.
- Permitirá mostrar la información de la institución, autoridades, docentes y estudiantes.

4.1.2. FACTIBILIDAD DEL PROYECTO

4.1.2.1. FACTIBILIDAD OPERATIVA

Se comparará el Sistema Actual con el Sistema Propuesto, debido a que la Institución posee un sistema manual, se tomará en cuenta ciertos parámetros necesarios para dar a conocer a la institución por medio de un Portal Semántico, para revisar si es operativa la solución propuesta.

A continuación se muestra una tabla con los parámetros a satisfacer en el Portal

Propuesto:

Parámetros	Sistema Actual	Sistema Propuesto
Ingreso de Datos	Microsoft Excel, Microsoft Office	Base de Datos
Organización de información y el conocimiento	No existe	Documento XTM
Presentación de la información y el conocimiento	Archivos físicos y digitales	Portal Semántico
Transparencia de la información	Interna	Externa
Usabilidad	Solo personal Autorizado	Utilizado por cualquier usuario
Navegabilidad	No existe	Fácil, ágil
Eficiencia	No existe	Búsqueda semántica basado en un tópico específico

Tabla IV.9. Parámetros a satisfacer por el portal

El sistema propuesto optimiza costo, recursos y tiempo frente al sistema con el que se trabaja actualmente; esto beneficia a la Institución directamente y a los usuarios (personal que va a interactuar directa o indirectamente con el sistema). Además, el sistema propuesto, crear un Portal para darlo a conocer en la web, utilizando una semántica adecuada para el usuario en la que podrá navegar de manera ágil, fácil y obteniendo los resultados deseados por lo tanto el desarrollo del sistema es factible operacionalmente.

4.1.2.2. FACTIBILIDAD TÉCNICA

Recursos requeridos para el portal:

- Recurso Hardware.
- Recurso Software.
- Recurso Humano.

RECURSO HARDWARE

Hardware Reutilizable

Para el desarrollo e implementación del Portal Semántico, la institución cuenta con el equipo y la infraestructura necesaria.

RECURSO SOFTWARE

Software Reutilizable

DESCRIPCION	SOFTWARE	VALOR
Sistema Operativo	Windows XP Professional	30,33
Documentación	Microsoft Office 2007 Professional	44.00
TOTAL		74.33

Tabla IV.10. Software Reutilizable

RECURSO HUMANO

Capacitación Recurso Humano

Para el sistema propuesto se necesita dar capacitación al Recurso Humano ya que al ser un Portal Semántico que posee una base de datos de la cual se abstraerá la información buscada, el personal deberá conocer sobre cómo llenar la base de datos “Liceo”.

4.1.2.3. FACTIBILIDAD ECONÓMICA

En las siguientes tablas se da a conocer los costos por inversión como por explotación del sistema actual y del sistema propuesto.

Sistema Actual

En la actualidad no se cuenta con un sistema que brinde la facilidad a los usuarios de obtener información relevante a la Institución de forma pública.

Sistema Propuesto

	INVERSIÓN		EXPLOTACIÓN	
	DESCRIPC.	COSTO	DESCRIPC.	COSTO MENSUAL
Implantación del Sistema	Costo total del proyecto	\$2766	Mantenimiento	\$50.00
TOTAL		\$2766		\$50.00

Tabla IV.11. Costo de Sistema Propuesto

BENEFICIOS**BENEFICIOS TANGIBLES**

- Mostrar información y conocimiento organizado de forma semántica.
- Darse a conocer en la web, teniendo la información pública.
- Actualización de información.

BENEFICIOS INTANGIBLES

- Acceso fácil y eficiente a la información.
- Satisfacción del personal y usuarios por la fácil e intuitiva navegación en el Portal Semántico.

4.1.2.4. FACTIBILIDAD LEGAL

Debido a que el Liceo Stephen Hawking es una institución de Educación, cuyo funcionamiento fue autorizado el 1 de julio de 2001, se rige por la Constitución Política del Estado Ecuatoriano, La CONESUP, el Estatuto Institucional y sus reglamentos, es decir; su funcionamiento se enmarca dentro del aspecto legal, es reconocida a nivel nacional y cuenta con mucho prestigio.

Antes de implantar el Portal Semántico de la Institución, se sugiere realizar una revisión del estatuto del Liceo Stephen Hawking, para que esté de acuerdo con el portal a desarrollarse, de esta manera el portal podrá ser realizado legalmente respetando las leyes de la Constitución de la República del Ecuador y las leyes establecidas por el Liceo Stephen Hawking.

Por estas razones el Sistema Propuesto es legalmente factible ya que no existe impedimento en que cada dependencia progrese para el bienestar de la Institución.

4.1.3. ESPECIFICACIÓN DE REQUERIMIENTOS DE SOFTWARE (SRS)**4.1.3.1. INTRODUCCIÓN**

En el presente documento presentamos las Especificaciones de Requerimiento de Software (SRS), lo cual permitirá establecer las bases para el desarrollo del

Portal Semántico “Liceo Stephen Hawking”. Con éste estudio minucioso podemos revelar posibles omisiones, malentendidos e inconsistencias que se pueden dar al principio del desarrollo, donde estos problemas son más fáciles de corregir.

A continuación este documento describe la información recopilada, en que nos basaremos para elaborar el Portal Semántico de la Institución.

PROPÓSITO

El objetivo de la especificación es definir de manera clara y precisa todas las funciones y restricciones del sistema que se desea construir. El documento está dirigido al desarrollador, al guía del proyecto y a los usuarios finales del sistema. Este documento será el instrumento de comunicación entre las partes implicadas, tomando parte en su confección miembros de cada una de las partes.

ÁMBITO DEL SISTEMA

En la actualidad no se cuenta con ninguna aplicación que represente la información y el conocimiento relevante de la Institución Liceo Stephen Hawking.

A partir de lo descrito en todo el documento presentamos un análisis de los problemas a los cuales vamos a buscar dar solución.

DEFINICIONES, ACRÓNIMOS Y ABREVIATURAS

DEFINICIONES

Usuarios	Personas que tienen acceso al portal
Secretaria	Persona que tiene la función de manejar la información tales como: Información Institucional Información de Autoridades Información de Estudiantes Información de Docentes
Técnico	Persona que se encarga de controlar el funcionamiento del Portal Semántico de la Institución u otros sistemas involucrados
Desarrolladora	Persona la cual ha realizado el portal

ACRÓNIMOS

SRS	Software Requirements Specification
XML	Extensible Markup Language

ABREVIATURAS

RR-HH	Recursos Humanos
XTM	XML Topic Maps

REFERENCIAS

IEEE Recommend Practice for Software Requirements Specification ANSI/IEEE (Estandar guía que hemos utilizado).

VISIÓN GENERAL DEL DOCUMENTO

Este documento se ha dividido en tres partes: La primera se refiere a una introducción que provee una versión general del SRS. La segunda da una visión general de la aplicación, con el objetivo de identificar las funciones que deben realizarse (con sus respectivas limitaciones). La tercera parte trata específicamente de los requerimientos de los usuarios a fin de satisfacer sus necesidades.

4.1.3.2. DESCRIPCIÓN GENERAL

En esta área se presenta una descripción de alto nivel del sistema. Se presentarán las funciones principales que deberá realizar el Sistema, así también como detallar sus principales restricciones

PERSPECTIVAS DEL PRODUCTO

El Portal Semántico de la Institución, en su primera versión estará interactuando con una base de datos: Base de Datos “Liceo” que contienen información de la Institución, Autoridades, Estudiantes, Docentes, Materias, Grados y Horarios.

FUNCIONES DEL SISTEMA

En una forma general, se detallará las tareas principales del portal:

- Representar información y conocimiento basado en el paradigma XML Topic Maps (XTM), creando un Portal Semántico el cual sea navegable y eficiente para los usuarios.

Ingreso de Usuarios al Portal Semántico

El sistema estará en la capacidad de mostrar información a todos los usuarios que deseen conocer sobre la Institución, de manera fácil, eficaz e intuitiva.

Gestión de representación de la Información y Conocimiento

Al momento de acceder al Portal Semántico, permitirá la navegación intuitiva.

Las tareas que realiza este módulo son:

Representación de la Información y Conocimiento: La representación se muestra en el Portal Semántico, teniendo procesos internos enfocados al documento XTM y a la base de datos “Liceo”, los mismos que facilitarán su comprensión e interpretación y harán de la navegación para el usuario más intuitiva y eficiente.

Creación de la Bases de datos.

De acuerdo a los datos recolectados, y necesidades de la Institución se creó una base de datos “Liceo”, con el fin de abstraer datos relevantes en el Portal Semántico y para futuros sistemas.

CARACTERÍSTICAS DE LOS USUARIOS

El Portal contará con interfaces amigables, fáciles de aprender y sencillas de manejar; presentado a los usuarios y autoridades una aplicación entendible para el manejo del mismo.

RESTRICCIONES

Dado que el sistema implementará la política y los reglamentos vigentes actuales que lo conforman es de esperar que futuros cambios en los modos de trabajo o en políticas, ejerzan un fuerte impacto sobre el sistema.

En cuanto a las restricciones Hardware/Software el sistema funcionara bajo el paradigma WEB.

DEPENDENCIAS

El Portal Semántico no tiene dependencias.

El sistema seguirá una arquitectura Web, por lo que la disponibilidad del sistema dependerá de la disponibilidad que tengan las PC's a Internet y el servidor en donde será alojada la Aplicación.

4.1.3.3. REQUISITOS ESPECÍFICOS

En esta tercera parte del SRS, se representan los requisitos funcionales que deberán ser satisfechos por el portal (todos los requerimientos son esenciales para un óptimo funcionamiento)

REQUISITOS FUNCIONALES

GESTIÓN DE INTEGRACIÓN DE BASE DE DATOS

REQ (1).- La base de datos deberá presentar un recuadro para logueo pidiendo usuario y contraseña.

REQ (2).- La base de datos deberá validar los datos ingresados y abrir la base de datos en caso de éxito, o no acceder en caso de fracaso.

REQ (3).- En caso de éxito los datos deberán llenarse y validarse antes de la presentación del portal a los usuarios.

GESTIÓN DE INFORMACIÓN SEMÁNTICA

REQ (4).- El portal presenta información relevante de la institución en la parte central del portal y al lado izquierdo la navegación de acuerdo al tema a escoger.

REQ (5).- El portal permitirá navegar al usuario de acuerdo a la semántica mostrada, dejándolo escoger un tema específico.

REQ (6).- El portal guarda el tema específico y lo lleva a buscar dentro del documento XTM.

REQ (7).- El portal encuentra la o las asociaciones para el tema escogido, y devuelve la información específica de acuerdo a la semántica presentada.

REQ (8).- El portal busca la información en la base de datos de acuerdo a la información solicitada y la muestra al usuario.

REQUISITOS DE INTERFACES EXTERNAS

INTERFACES DE USUARIO

La interfaz de usuario va a ser clara, precisa, manejará una interfaz amigable que contemple colores sobrios, permita una navegación adecuada en la web, la cual contendrá accesos correctos y controlados, el manejo de la aplicación se realizara a través de teclado y mouse.

INTERFACES DE COMUNICACIONES

La conexión a la red se establecerá por medio de una conexión Internet, se puede utilizar también conexiones vía intranet.

REQUISITOS DE RENDIMIENTO

El número de puestos a los que se debe dar servicio no es determinado debido a que estará conectada a Internet e intranet del Liceo Stephen Hawking, sin embargo el tiempo de respuesta en las operaciones debe ser menor a 10 s., dependiendo del número de clientes.

REQUISITOS DE DESARROLLO

La metodología utilizada para el desarrollo del producto es XP (Programación Extrema), nos permitirá trabajar con un equipo de desarrollo pequeño y garantizará el trabajo del desarrollador.

REQUISITOS TECNOLÓGICOS

La aplicación cliente se ejecutará sobre un PC con una configuración:

- Intel Pentium IV de 2.6 GHz
- Memoria: 1024 MB de memoria RAM
- Disco Duro: Samsung 120 GB
- Monitor digital SyncMaster 17"
- Tarjeta de video
- CD Writer
- Memoria CACHE 512 Kb
- Teclado y Mouse

4.1.3.4. ATRIBUTOS

SEGURIDAD

Cuando un usuario intenta conectarse al Portal Semántico de la Institución no será necesario que se loguee puesto que la información mostrada es pública.

Los datos que se ingresarán en la base de datos únicamente lo podrá alterar la Secretaria, Directora o Desarrollador, para lo cual se proveerá de un usuario y una clave única para el ingreso a ésta.

4.2 FASE DE DISEÑO

4.2.1. PRESUPUESTO DEL COSTO POR EL ESFUERZO DE DESARROLLO

Identificar una técnica de Estimación

Es un factor muy importante el tamaño y la complejidad del proyecto para poder escoger la técnica de estimación adecuada. Nuestro proyecto tiene una complejidad media, y su tamaño es relativamente pequeño, razones por las cuales hemos escogido los métodos de PUNTOS DE FUNCIÓN Y COCOMO II, el primero nos ayudará a identificar todos los requerimientos con sus respectivas complejidades, y la herramienta COCOMO II nos permitirá encontrar el costo y la duración del proyecto lo utilizaremos a través de la herramienta y de forma manual.

4.2.1.1. COCOMO II

Total de líneas de código= 4736 líneas de código.

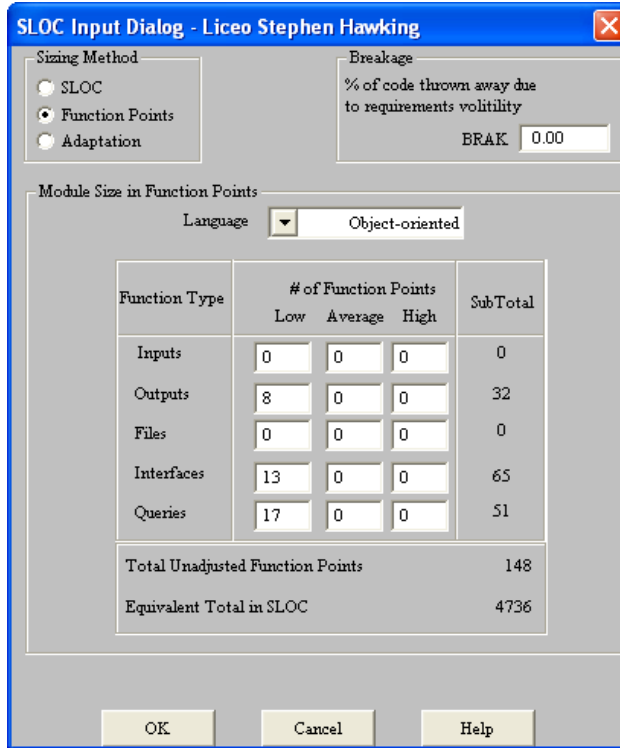


Figura IV.2. Software Cocomo

Total Costo de Desarrollo del Proyecto

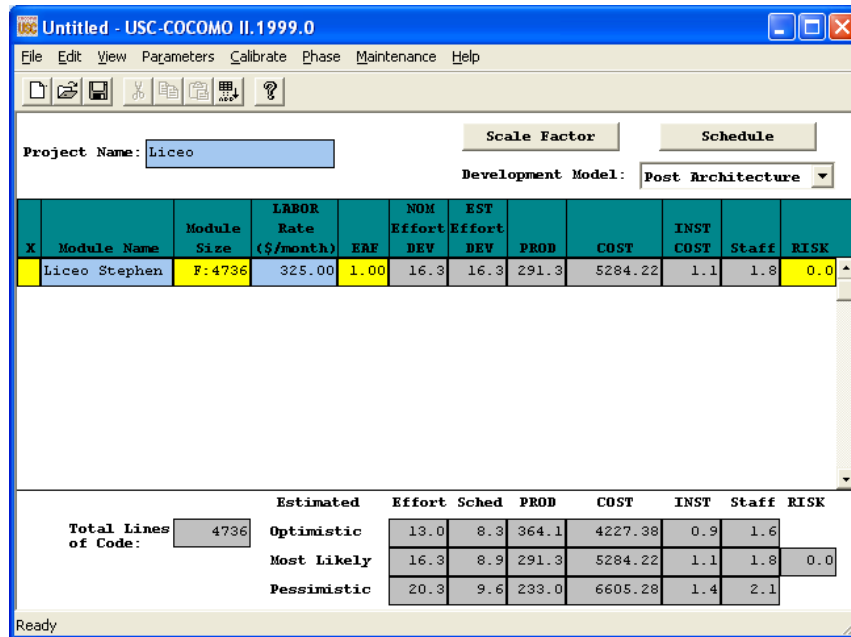


Figura IV.3. Total costo de desarrollo según Cocomo

Total de persona mes a desarrollar un proyecto.

Phase Distribution - Project Plans & Requirements				
=====				
Life Cycle Phase	Plans And Requirements			
Life Cycle Effort	1.138 Person Months			
Life Cycle Schedule	1.506 Months			
=====				
	PCNT	EFFORT (PM)	SCHEDULE	Staff
Requirements Analysis	47.544	0.541	1.506	0.359
Product Design	16.228	0.185	1.506	0.123
Programming	2.956	0.034	1.506	0.022
Test Planning	2.728	0.031	1.506	0.021
Verification and Validation	6.228	0.071	1.506	0.047
Project Office	15.044	0.171	1.506	0.114
CM/QA	3.272	0.037	1.506	0.025
Manuals	6.000	0.068	1.506	0.045

OK Help

Figura IV.4. Total planes y requerimientos según Cocomo

Tiempo a desarrollar y número de personas para la Fase de Ingeniería de Requerimiento.

Phase Distribution - Project Product Design				
=====				
Life Cycle Phase	Product Design			
Life Cycle Effort	2.764 Person Months			
Life Cycle Schedule	2.178 Months			
=====				
	PCNT	EFFORT (PM)	SCHEDULE	Staff
Requirements Analysis	12.500	0.346	2.178	0.159
Product Design	41.000	1.133	2.178	0.520
Programming	12.228	0.338	2.178	0.155
Test Planning	4.728	0.131	2.178	0.060
Verification and Validation	6.228	0.172	2.178	0.079
Project Office	12.544	0.347	2.178	0.159
CM/QA	2.772	0.077	2.178	0.035
Manuals	8.000	0.221	2.178	0.102

OK Help

Figura IV.5. Total diseño del producto según Cocomo

Tiempo a desarrollar y número de personas para la Fase de Diseño

Phase Distribution - Project Programming				
Life Cycle Phase	Programming			
Life Cycle Effort	10.183 Person Months			
Life Cycle Schedule	4.825 Months			
	PCNT	EFFORT (PM)	SCHEDULE	Staff
Requirements Analysis	4.000	0.407	4.825	0.084
Product Design	8.000	0.815	4.825	0.169
Programming	56.500	5.754	4.825	1.192
Test Planning	4.228	0.431	4.825	0.089
Verification and Validation	7.228	0.736	4.825	0.153
Project Office	7.272	0.741	4.825	0.153
CM/QA	6.772	0.690	4.825	0.143
Manuals	6.000	0.611	4.825	0.127

Figura IV.6. Total Programación del proyecto según Cocomo

Tiempo a desarrollar y número de personas para la Fase de Programación.

Phase Distribution - Project Integration & Test				
Life Cycle Phase	Integration and Test			
Life Cycle Effort	3.312 Person Months			
Life Cycle Schedule	1.903 Months			
	PCNT	EFFORT (PM)	SCHEDULE	Staff
Requirements Analysis	2.500	0.083	1.903	0.044
Product Design	5.000	0.166	1.903	0.087
Programming	33.912	1.123	1.903	0.590
Test Planning	2.500	0.083	1.903	0.044
Verification and Validation	31.544	1.045	1.903	0.549
Project Office	8.272	0.274	1.903	0.144
CM/QA	8.272	0.274	1.903	0.144
Manuals	8.000	0.265	1.903	0.139

Figura IV.7. Total Pruebas e Integración según Cocomo

Tiempo a desarrollar y número de personas para la Fase de Integración y Pruebas.

PUNTOS DE FUNCIÓN

Parámetro de Medición	Factor de ponderación				PF
	Cuentas	Simple	Medio	Complejo	Parciales
Número de entradas de usuario	4	3	<u>4</u>	6	16
Número de salidas de	10	4	5	<u>7</u>	70

usuario					
Número de peticiones de usuario	20	3	34	<u>6</u>	120
Número de archivo	3	<u>7</u>	10	15	21
Número de Interfaces externas	4	5	<u>7</u>	11	28
Cuenta Total					255

Tabla IV.12. Total Puntos de Función

Para tomar los valores de pesimista y optimista en los puntos de función, se sumará y restará el valor de 3 al valor probable.

Factor de Costo



Figura IV.8. Escala de Complejidad

Valor de Dominio de Puntos de Función	Puntos de Función					PF
	Optimista	Probable	Pesimista	Cuenta Estimada	Peso	Ctas
Número de entradas de usuario	13	16	19	16	4	64
Número de salidas de usuario	67	70	73	70	5	350
Número de peticiones de usuario	117	120	123	120	4	480
Número de archivo	18	21	24	21	7	147
Número de Interfaces externas	25	28	31	28	6	168
Cuenta Total						1209

Tabla IV.13. Factor costo Puntos de Función

A continuación se estimara los factores de Ponderación de Complejidad

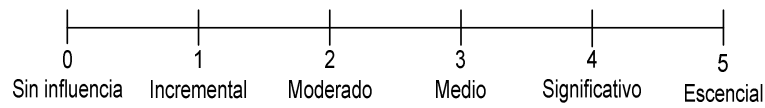


Figura IV.9. Escala de ponderación

FACTOR	VALOR
Copia de seguridad recuperación	4
Comunicación de datos	4
Replica de Datos	5
Consumo de Servicios Web	3
Rendimiento crítico	2
Entorno operativo existente	3
Entrada de datos en línea	4
Transacciones de entrada en múltiples pantallas	3
Archivos maestros actualizados en línea	4
Complejidad de valores del dominio de información	4
Complejidad del procesamiento interno	4
Código diseñado para la reutilización	5
Conversión/instalación en diseño	2
Instalaciones múltiples	1
Aplicación diseñada para el cambio	3
Complejidad del diseño de la base de Datos	3
Complejidad Diseño de Interfaces de la Aplicación	3
Factor de ajuste de la complejidad	3.35

Tabla IV.14. Factores de ponderación de complejidad

Finalmente, con los datos obtenidos de las tablas anteriores se definirá el número estimado de puntos de función:

$$\text{Cuenta Total} = 1209$$

$$\sum F_i = 3.35$$

$$PF_{\text{ESTIMADO}} = \text{CUENTA TOTAL} \times (0,65 + 0,01 + \sum F_i)$$

$$PF_{\text{ESTIMADO}} = 1209 \times (0,65 + 0,01 + 3.35)$$

$$PF_{\text{ESTIMADO}} = 3.849$$

El modelo de estimación que utilizaremos será el Modelo de COCOMO (Modelo Constructivo de Costo) orgánico.

Proyecto de Software	a_b	b_b	c_b	d_b
Orgánico	2,4	1,05	2,5	0,38
Semiacoplado	3,0	1,12	2,5	0,35
Empotrado	3,6	1,20	2,5	0,32

Tabla IV.15. Modelo Constructivo de Costo según Cocomo

$$KPF = 1,6$$

$$E = a_i KPF^{b_i}$$

$$E = 2,4(1,6)^{1,05}$$

$$E = 3,9 \text{ personas} - \text{mes}$$

Analizando la respuesta obtenida se necesitarán para el proyecto 4 personas.

$$D = (c_b)(E^{d_b})$$

$$D = (2,5)(4^{0,38})$$

$$D = 4 \text{ meses}$$

La programación tendrá una duración aproximada de 4 meses.

4.2.1.2. PRESUPUESTOS DE LOS RECURSOS DE ENTORNO

CARACTERIZACIÓN DEL RECURSO HUMANO

Programadores:

A continuación sacaremos el costo con un valor de \$1,50 la hora, por 7 días a la semana con un trabajo de 8 horas, en un mes.

RECURSO	CANT.	COSTO	SUBTOTAL
Programadores	1	336,00	336,00
TOTAL			\$ 336,00

Tabla IV.16. Costo de trabajador por mes

CARACTERIZACIÓN DEL HARDWARE

RECURSO	CANT.	COSTO	SUBTOTAL
Impresora	1	120,00	120,00
Puntos de Red	1	30,00	30,00
PC Portátil HP Intel, 1Gb de Ram, 80 GB	1	1.100,00	1.100,00
TOTAL			\$1.250,00

*Tabla IV.17 Caracterización del Hardware***CARACTERIZACIÓN DEL SOFTWARE PARA EL DESARROLLO**

DESCRIPCIÓN	SOFTWARE	VALOR
Sistema Operativo	Windows XP	30,00
Lenguaje de Programación	Microsoft Visual Studio.NET 2008	40,00
Manejador de Bases de Datos	SQL Server 2005	40,00
Documentación	Office 2007 Professional	30,00
Documentación	Project 2000	25,00
Herramienta Colaborativa	XP	300,00
TOTAL		\$ 465,00

*Tabla IV.18. Caracterización del Software***CARACTERIZACIÓN DEL SOFTWARE PARA LA IMPLANTACIÓN**

DESCRIPCION	SOFTWARE	VALOR
Sistema Operativo para la Aplicación Cliente y Servidor	Windows XP	30,00
Manejador de Bases de Datos	SQL Server 2005	40,00
Lenguaje de Programación	Microsoft Visual Studio.NET 2008	40,00
TOTAL		\$ 110,00

*Tabla IV.19. Caracterización del software para la implantación***COSTO DEL SOFTWARE**

DESCRIPCION	VALOR (\$)
Líneas de Códigos (LDC)	1.209,00

Pago de Sueldos a los Desarrolladores por tres meses y medio.	1.176,00
Depreciaciones de los Equipos (33% anual)	676,50
Depreciación Por Utilización De Las Licencias Del Software	189,75
TOTAL	3251,25

Tabla IV.20. Costo del Software

COSTO TOTAL DEL PROYECTO

DESCRIPCION	VALOR (\$)
Valor del software	2.385,00
Software para la implantación	465,00
Hardware para la implantación	1.250,00
Mobiliario	60,00
Subtotal	4.160,00
IVA (12%)	499,20
TOTAL	4.659,20

Tabla IV.21. Costo Total del Proyecto

CONCLUSION:

PROYECTO	COSTO
HERRAMIENTA COCOMO	\$5.284,22
MANULAMENTE	\$4.659,20

Tabla IV.22. Conclusión de costos del Proyecto

Hemos escogido el cálculo de las líneas de punto de fusión realizadas manualmente por contener un costo menor **\$4.659,20**.

4.2.1.3. ESTIMACIONES DE ESFUERZO BASADAS EN HISTORIAS DE USUARIOS

Las estimaciones de esfuerzo asociado a la implementación de las historias se establecen utilizando como medida el punto. Un punto, equivale a una semana ideal de programación. Las historias generalmente valen de 1 a 3 puntos. Por

otra parte, se mantiene un registro de la “velocidad” de desarrollo, establecida en puntos por iteración, basándose principalmente en la suma de puntos correspondientes a las historias de usuario que fueron terminadas en la fase de exploración.

MÓDULO DE INTEGRACIÓN DE BASE DE DATOS

Ingreso de datos 2

Procesos de búsqueda 2

MÓDULO DE BUSQUEDA SEMÁNTICA DE INFORMACIÓN

Creación de documento XTM 3

Búsqueda Semántica 3

Gestión de Requisitos 2

MÓDULO DE FUNCIONALIDAD DEL USUARIO

Mostrar información 2

Consulta de información de requisitos 1

4.2.2. PLANIFICACIÓN Y SEGUIMIENTO DEL PROYECTO

Partiendo de las historias de usuario anteriores se realiza una planificación en cuatro iteraciones basándose en el tiempo y procurando agrupar la funcionalidad relacionada en la misma iteración.

4.2.2.1 FECHAS DE ENTREGA

Nombre de tarea	Duración	Comienzo	Fin	Predecesoras	Nombres de los recursos
1 INICIO	1 día	lun 01/06/09	lun 01/06/09		
2 CAPITULO I: Marco Referencial	1 día	mar 02/06/09	mar 02/06/09	1	
3 Antecedentes	1 día	mar 02/06/09	mar 02/06/09		Desarrolladora
4 Justificación	1 día	mar 02/06/09	mar 02/06/09	3CC	Desarrolladora
5 Objetivo General	1 día	mar 02/06/09	mar 02/06/09	4CC	Desarrolladora
6 Objetivos Especificos	1 día	mar 02/06/09	mar 02/06/09	5CC	Desarrolladora
7 Hipotesis	1 día	mar 02/06/09	mar 02/06/09	6CC	Desarrolladora
8 CAPITULO II: Marco Teorico	27 días?	mié 03/06/09	jue 09/07/09	7	
9 Representación del conocimiento	9 días	mié 03/06/09	lun 15/06/09		
10 Ontologías	3 días	mié 03/06/09	vie 05/06/09		Desarrolladora
11 Tesoros	3 días	lun 08/06/09	mié 10/06/09	10	Desarrolladora
12 Topic Maps	3 días	jue 11/06/09	lun 15/06/09	11	Desarrolladora
13 Representación de la información en la web	8 días?	mar 16/06/09	jue 25/06/09	12	
14 SGML	1 día	mar 16/06/09	mar 16/06/09		Desarrolladora
15 Tecnología HTML	5 días	mié 17/06/09	mar 23/06/09	14	Desarrolladora
16 Tecnología XHTML	1 día?	mié 24/06/09	mié 24/06/09	15	Desarrolladora
17 Tecnología XML	1 día?	jue 25/06/09	jue 25/06/09	16	Desarrolladora
18 Representación del conocimiento bajo Topic Maps TM	10 días?	vie 26/06/09	jue 09/07/09	17	
19 Definición de TM	1 día?	vie 26/06/09	vie 26/06/09		Desarrolladora
20 Evolución de TM	1 día?	lun 29/06/09	lun 29/06/09	19	Desarrolladora
21 Modelos de TM	4 días	mar 30/06/09	vie 03/07/09	20	
22 TMGL	1 día	mar 30/06/09	mar 30/06/09		Desarrolladora
23 TMCL	1 día	mié 01/07/09	mié 01/07/09	22	Desarrolladora
24 XTM	1 día	jue 02/07/09	jue 02/07/09	23	Desarrolladora
25 HyTM	1 día	vie 03/07/09	vie 03/07/09	24	Desarrolladora
26 Funcionamiento	2 días	lun 06/07/09	mar 07/07/09	25	Desarrolladora
27 Ventajas y desventajas	2 días	mié 08/07/09	jue 09/07/09	26	Desarrolladora
28 CAPITULO III: Representación del conocimiento bajo XML Topic Maps XTM	25 días	vie 10/07/09	jue 13/08/09	27	
29 Definición de XTM	3 días	vie 10/07/09	mar 14/07/09		Desarrolladora
30 Evolución de XTM	3 días	mié 15/07/09	vie 17/07/09	29	Desarrolladora
31 Funcionamiento de XTM	5 días	lun 20/07/09	vie 24/07/09	30	Desarrolladora
32 Ventajas y desventajas de XTM	3 días	lun 27/07/09	mié 29/07/09	31	Desarrolladora
33 Normativas actuales	3 días	jue 30/07/09	lun 03/08/09	32	Desarrolladora
34 Software	8 días	mar 04/08/09	jue 13/08/09	33	Desarrolladora
35 CAPITULO IV: Desarrollo e implementación del portal semántico	54 días	vie 14/08/09	mié 28/10/09	34	
36 FASE DE ANALISIS	9 días	vie 14/08/09	mié 26/08/09		
37 Requerimentación Inicial	3 días	vie 14/08/09	mar 18/08/09		Desarrolladora
38 Especificación de Requerimientos de Software (SRS)	3 días	mié 19/08/09	vie 21/08/09	37	Desarrolladora
39 Analisis	3 días	lun 24/08/09	mié 26/08/09	38	Desarrolladora
40 FASE DE DISEÑO	10 días	jue 27/08/09	mié 09/09/09	39	
41 Diseño	5 días	jue 27/08/09	mié 02/09/09		Desarrolladora
42 Diagramas	5 días	jue 03/09/09	mié 09/09/09	41	Desarrolladora
43 FASE DE IMPLEMENTACION	35 días	jue 10/09/09	mié 28/10/09	42	
44 Implementación	30 días	jue 10/09/09	mié 21/10/09		Desarrolladora
45 Pruebas	5 días	jue 22/10/09	mié 28/10/09	44	Desarrolladora
46 FIN	0 días	mié 28/10/09	mié 28/10/09	45	Desarrolladora

Figura IV.10. Fechas de entrega del proyecto

DIAGRAMAS GANTT

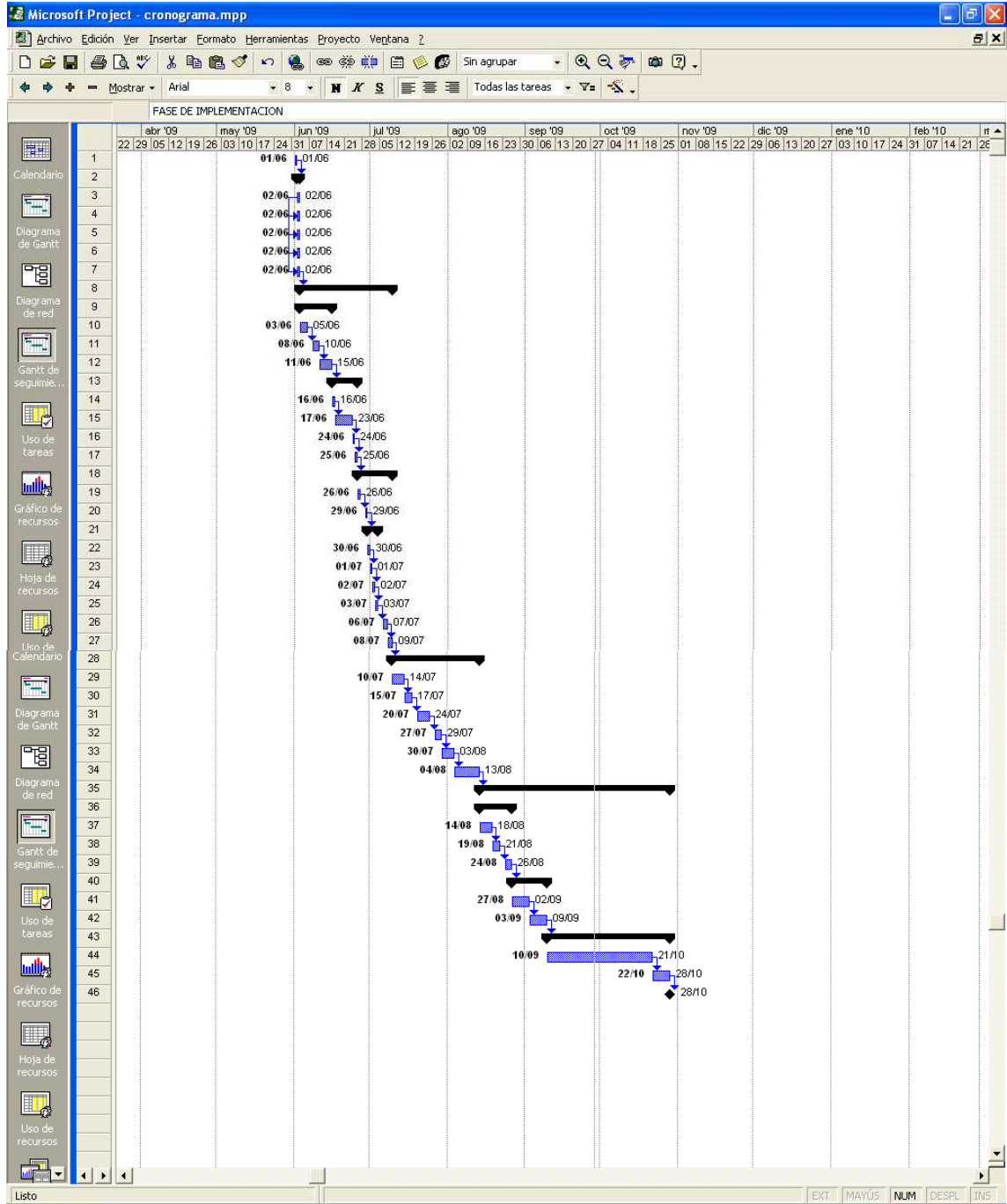


Figura IV.11. Diagrama Gantt

4.2.2.2 PRIMERA ITERACIÓN: Prototipo

En esta primera iteración se creará un prototipo con el que se comprobará la adecuación de la tecnología escogida y se intentará crear la mayor parte de la base de la arquitectura del sistema. La funcionalidad no es muy extensa tan sólo un mínimo para tener cuanto antes una versión demo que poder mostrar en el sitio Web y así atraer posibles usuarios.

Historia primera iteración

FUNCIONALIDAD COMÚN	2
ESTIMACIÓN INICIAL	2

4.2.2.3 SEGUNDA ITERACIÓN: Carga de datos

En una segunda iteración se añadirá la funcionalidad necesaria para gestionar la carga de datos a la base de datos “Liceo”.

La duración real de la iteración ha sido muy breve gracias a que el núcleo del sistema ya había sido realizado en la iteración anterior con todo el esfuerzo de integración de tecnologías. También se hará que el sistema sea accesible desde cualquier PC con conexión a Internet.

Historias segunda iteración

CARGA DE DATOS	2
CONTROL DE ACTUALIZACIÓN DE LAS BASES DE DATOS	1
ACCESIBILIDAD DESDE INTERNET	1
ESTIMACIÓN INICIAL	4
REAL	4

4.2.2.4 TERCERA ITERACIÓN: Módulo Información

Semántica

En esta tercera iteración se añadirá la funcionalidad relativa a la integración del documento XTM al Portal. Una vez hecho esto, la gestión de búsqueda semántica se ha procedido a desarrollar e implementar el Portal Semántico.

Historias tercera iteración

CREACIÓN DEL DOCUMENTO XTM	2
----------------------------	---

PROCESAMIENTO SEMÁNTICO	2
ESTIMACION INICIAL	4
REAL	4

4.2.2.5 CUARTA ITERACIÓN: Módulo de Funcionalidad del Usuario

La cuarta iteración conllevará la finalización del sistema tras el procesamiento de la información que contienen las bases de datos y la posterior generación de búsquedas semánticas a través de la creación del documento XTM para la representación de la información y el conocimiento.

Historias cuarta iteración

PROCESAMIENTO DE LA INFORMACIÓN	2
PROCESAMIENTO DEL CONOCIMIENTO	1
ESTIMACIÓN INICIAL	3
REAL	3

4.2.3. GESTIÓN DE RIESGOS.

Los riesgos se pueden clasificar en diferentes grupos para poder identificarlos basándonos en un estándar para realizar dichas especificaciones.

4.2.3.1 Riesgos Genéricos

Riesgo	Tipo Riesgo	Descripción
Problemas financieros de la organización.	Proyecto	Problemas de escaso presupuesto
Desintegración del equipo de trabajo	Proyecto	Los integrantes del grupo salgan antes de que el proyecto sea completado.

Hardware no disponible	Proyecto	Hardware que es indispensable para la elaboración no sea entregado a tiempo
Cambio de Requerimientos	Proyecto y producto	Habr� cambios a los requerimientos inicialmente definidos.
Retrasos en la entrega	Proyecto y producto	El producto no este disponible a tiempo
Falta de Planificaci�n	Proyecto	El sistema no cumpla con el cronograma de actividades establecidas.
Cortes de Energ�a el�ctrica	Proyecto	Por estas situaciones hay retrasos en el avance, esto puede conllevar a no entregar el proyecto la fecha establecida.
Cambio de tecnolog�a	Instituci�n	La tecnolog�a en que el sistema es construido es reemplazada por nueva tecnolog�a
Competencia del producto	Instituci�n	Un producto competitivo est� publicitado antes de que el sistema est� completo
Impuntualidad del Personal	Proyecto	Provoca retrasos en las actividades establecidas.
Enfermedad del Personal	Proyecto	Las tareas asignadas a cada integrante del equipo no sean terminadas a tiempo esto retrasa el avance.
Robo del Disco Duro o Hardware indispensable	Proyecto	Conllea a que el sistema no sea terminado a tiempo o definitivamente no se termine puesto que toda la informaci�n

		estuvo centralizada en este.
--	--	------------------------------

Tabla IV.23. Riesgos Genéricos

4.2.3.2 Categorización del Riesgo

Código	Efectos
1	Catastrófico
2	Serio
3	Tolerable

Tabla IV.24. Códigos y efectos para categorizar el riesgo

Código Riesgo	Probabilidad de que ocurre	Código de Riesgo consecuencia	Estimación Impacto	Documentos
01	75%	01	proy. = 9 Prod = 8	Doc 1
02	40%	02	proy. = 7 Prod = 6	Doc 2
03	40%	02	proy. = 7 Prod = 7	Doc 3
04	30%	02	proy. = 8 Prod = 8	Doc 4
05	20%	02	proy. = 7 Prod = 7	Doc 5
06	15%	02	Proy = 7 Prod = 6	Doc 6
07	15%	03	Proy = 8 Prod = 6	Doc 7
08	20%	01	Proy = 7 Prod = 6	Doc 8
09	15%	01	Proy = 6 Prod = 5	Doc 9
10	25%	03	Proy = 7 Prod = 4	Doc 10

11	25%	02	Proy = 8 Prod = 5	Doc 11
12	30%	02	proy. = 6 Prod = 5	Doc 12

Tabla IV.25. Categorización del riesgo

Después de haber analizado el impacto y la probabilidad de que los riesgos puedan ocurrir se ha establecido la información necesaria para la supervisión de cada uno de los riesgos, los cuales se basan en una documentación adicional estandarizada en donde se muestran las causas y el plan de aversión para contrarrestar los posibles riesgos que se presten en el desarrollo del proyecto.

4.2.3.3. Plan de Contingencia del Riesgo

Doc1

IDENTIFICACIÓN

Problemas financieros de la organización

CAUSAS

Costos cambiantes en el mercado.

Nuevos requerimientos a lo largo del desarrollo del proyecto.

PLAN DE AVERSIÓN

Preparar un informe para los directivos de la institución que muestre de que manera el proyecto va a aportar al buen desenvolvimiento de las actividades de la institución. De este modo estamos asegurando la aprobación del presupuesto.

RESPONSABLE

Departamento Financiero

Doc2

IDENTIFICACIÓN

Desintegración del equipo de trabajo

CAUSAS

Poca experiencia.

Falta de organización.

Mala distribución del tiempo disponible

PLAN DE AVERSIÓN

Organizar y distribuir adecuadamente el plazo de tiempo disponible para terminar el proyecto, utilizando agendas.

Realizar los respectivos contratos de trabajo, considerando cláusulas de incumplimiento, de manera que no se pierda económicamente, además tener una persona ya lista para que inmediatamente pueda cubrir esa falta y no retrase la elaboración del sistema.

RESPONSABLE

Pamela Fuertes

Doc3

IDENTIFICACIÓN

Hardware no disponible.

CAUSAS

Falta de recursos económicos

Daños imprevistos o falta de HW

Equipos con software sin licencias

PLAN DE AVERSIÓN

Notificar a los Directivos de la Institución que no se cuenta con el hardware necesario y la implantación del sistema no se la puede realizar,

RESPONSABLE

Directivos del Liceo Stephen Hawking

Pamela Fuertes

Doc4

IDENTIFICACIÓN

Cambio de Requerimientos

CAUSAS

Poca experiencia.

Mala recolección de la información.

PLAN DE AVERSIÓN

Lograr una buena comunicación con el cliente.

Recolectar información correcta utilizando técnicas adecuadas para lograr la mayor captación.

RESPONSABLE

Pamela Fuertes

Doc5

IDENTIFICACIÓN

Retrasos en la entrega

CAUSAS

Elaboración del cronograma de planificación no estuvo bien realizada.

Irresponsabilidad del personal

PLAN DE AVERSIÓN

Considerar todas las actividades al momento de planificar el desarrollo del sistema

Establecer al inicio las condiciones para llevar a cabo el desarrollo del sistema puede ser definir días y horas de trabajo.

RESPONSABLE

Pamela Fuertes

Doc6

IDENTIFICACIÓN

Falta de Planificación

CAUSAS

Falta de experiencia

Falta de conocimiento de herramientas de planificación

PLAN DE AVERSIÓN

Dar capacitación para hacer uso de herramientas de planificación.

RESPONSABLE

Pamela Fuertes

Doc7

IDENTIFICACIÓN

Cortes de Energía eléctrica

CAUSAS

Desastres Naturales (Lluvias, rayos, temblores, etc.)

Fallas técnicas en las instalaciones de la institución como en la central de distribución.

PLAN DE AVERSIÓN

Contar un dispositivo regulador de voltaje ya que si hay fallas en el fluido eléctrico este puede causar daños en los equipos.

RESPONSABLE

Pamela Fuertes.

Doc8

IDENTIFICACIÓN

Cambio de tecnología

CAUSAS

Costos muy altos de licencias

Falta de recursos económicos

PLAN DE AVERSIÓN

Establecer las condiciones en las que el sistema se va a ejecutar con la finalidad de evitar estos contratiempos

RESPONSABLE

Pamela Fuertes

Doc9

IDENTIFICACIÓN

Impuntualidad del Personal

CAUSAS

Calamidad doméstica

Desinterés de la persona

PLAN DE AVERSIÓN

Organizar y distribuir adecuadamente tiempo disponible para terminar el proyecto

Realizar los respectivos contratos de trabajo, considerando cláusulas de incumplimiento.

RESPONSABLE

Pamela Fuertes

Doc10

IDENTIFICACIÓN

Enfermedad del Personal

CAUSAS

Fuera del alcance de la organización.

PLAN DE AVERSIÓN

Reorganizar el trabajo y crear documentos, para que al integrarse la persona nuevamente al equipo, pueda ponerse al corriente inmediatamente.

RESPONSABLE

Pamela Fuertes

Doc11

IDENTIFICACIÓN

Robo del Disco Duro o Hardware indispensable

CAUSAS

Falta de seguridad

PLAN DE AVERSIÓN

Tener los respectivos respaldos para que en caso de que ocurra poder avanzar con la ejecución del sistema y no retrasar su entrega.

RESPONSABLE

Pamela Fuertes

Doc12

IDENTIFICACIÓN

Organizacional

CAUSAS

Fuera del alcance de la organización.

PLAN DE AVERSIÓN

Notificar sobre el particular a los Directivos del Liceo Stephen Hawking ya que esta situación puede retrasar la entrega del sistema.

Contar con la información personal del Equipo de trabajo

RESPONSABLE

Pamela Fuertes

4.2.4. REFINAMIENTO DE LAS HISTORIAS DE USUARIOS

SISTEMA PROPUESTO

MÓDULO DE INTEGRACIÓN DE BASE DE DATOS

HISTORIA DE USUARIO	
Número: 1	Usuario: Directora, Secretaria.
Nombre historia: Base de Datos	
Prioridad en negocio: Alta	Riesgo en desarrollo: 04, 07, 09.
Puntos estimados: 8	Iteración asignada: 2
Programador responsable: Pamela Fuertes	
Descripción: Es necesario crear la Base de Datos “Liceo”; con la finalidad de automatizar los datos existentes en la Institución y que la información presentada se genere siempre con información actualizada.	
Observaciones: Es importante la creación de la Base de Datos en la Institución para facilitar el manejo de información, y para la integración de futuros sistemas	

Tabla IV.26. Refinamiento de la Historia de usuario 1

MÓDULO DE BÚSQUEDA SEMÁNTICA DE INFORMACIÓN

HISTORIA DE USUARIO	
Número: 2	Usuario: Cualquier usuario
Nombre historia: Búsqueda Semántica	
Prioridad en negocio: Alta	Riesgo en desarrollo: 03, 06, 07, 08, 12
Puntos estimados: 3	Iteración asignada: 3
Programador responsable: Pamela Fuertes	
<p>Descripción:</p> <p>Una vez que el usuario ingresa en el Portal Semántico, internamente el portal hace su búsqueda de acuerdo al tema escogido en el documento XML Topic Maps (XTM), y devuelve el o los temas a los que está directamente relacionado.</p>	
<p>Observaciones:</p> <p>Al momento de realizar la búsqueda semántica en el documento XTM, ayuda al usuario en la navegabilidad y es eficiente al mostrar información que proveerá la base de datos.</p>	

Tabla IV.27. Refinamiento de la Historia de usuario 2

MÓDULO DE FUNCIONALIDAD DEL USUARIO

HISTORIA DE USUARIO	
Número:3	Usuario: Cualquier usuario
Nombre historia: Información General de la Institución	
Prioridad en negocio: Alta	Riesgo en desarrollo: 02, 03, 04, 05, 08, 10, 11, 12
Puntos estimados: 5	Iteración asignada: 4
Programador responsable: Pamela Fuertes	
<p>Descripción:</p> <p>El usuario ingresa al Portal Semántico y accede a la información general de la institución como es la misión, visión, objetivos, historia, generalidades, autoridades, lista de docentes contratados, y pensum académico.</p>	
<p>Observaciones:</p> <p>Los usuarios que pueden acceder a esta información es cualquier persona interesada en conocer más sobre la Institución.</p> <p>El usuario notará que la navegabilidad en el Portal Semántico es sencilla y eficiente.</p>	

Tabla IV.28. Refinamiento de la Historia de usuario 3

4.2.4.1. REPRESENTACIÓN DE LAS HISTORIAS DE USUARIOS

MÓDULO DE INTEGRACIÓN DE BASE DE DATOS

Historia de usuario1

Descripción de las Interacciones:

ACTORES	SISTEMA
1. Los Usuarios (Directora, Secretaria) inicia la interacción ingresando a la base de datos.	2. Presenta Recuadro de ingreso de usuario y contraseña.

3. Usuario inicia sesión ingresando los datos solicitados.	4. Valida los datos y se conecta
5. Escoge la base de datos correcta e ingresa los datos	6. Valida datos y los almacena
CURSOS ALTERNATIVOS	
4,6: DATOS INCORRECTOS: Error de datos ingresados	

Tabla IV.29. Representación de la historia de usuario 1

REPRESENTACION GRAFICA

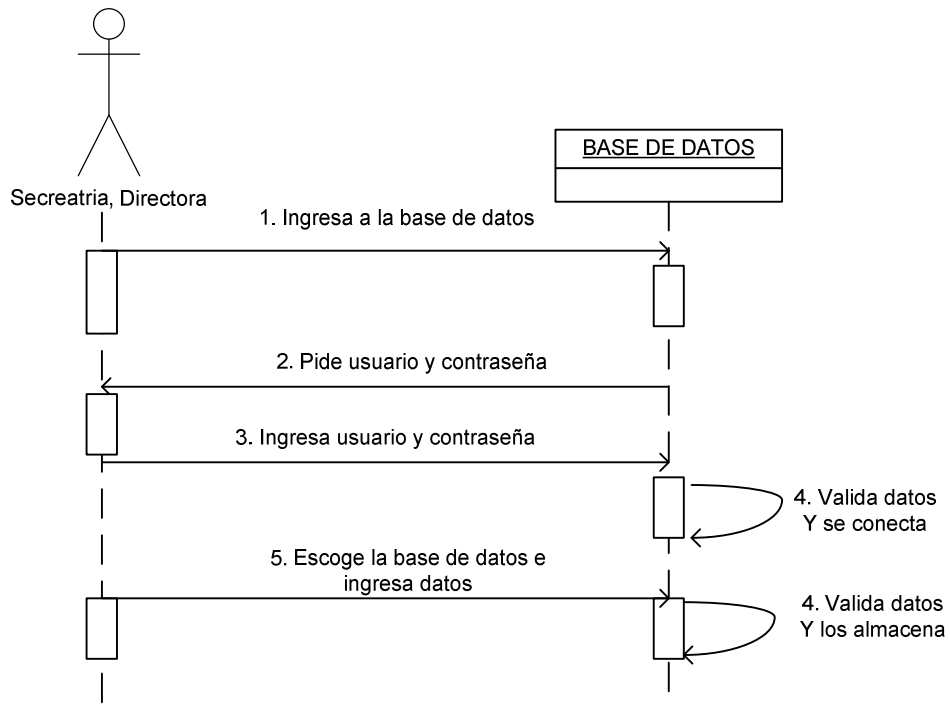


Figura IV.12. Representación grafica de la historia de usuario 1

MÓDULO DE BÚSQUEDA SEMÁNTICA DE INFORMACIÓN

Historia de usuario 2

Descripción de las Interacciones:

ACTORES	SISTEMA
1. Usuario ingresa a la página principal de la institución	2. Presenta información relevante de la institución
3. Navega por el portal, escogiendo un tema específico	4. Toma el tópico específico y lo busca dentro del documento XTM
	5. Encuentra la o las asociaciones a este tópico y devuelve la información específica a este tema de forma semántica
	6. Muestra la información que se encuentra en la base de datos de acuerdo a la semántica correspondiente

Tabla IV.30. Representación de la historia de usuario2

REPRESENTACION GRAFICA

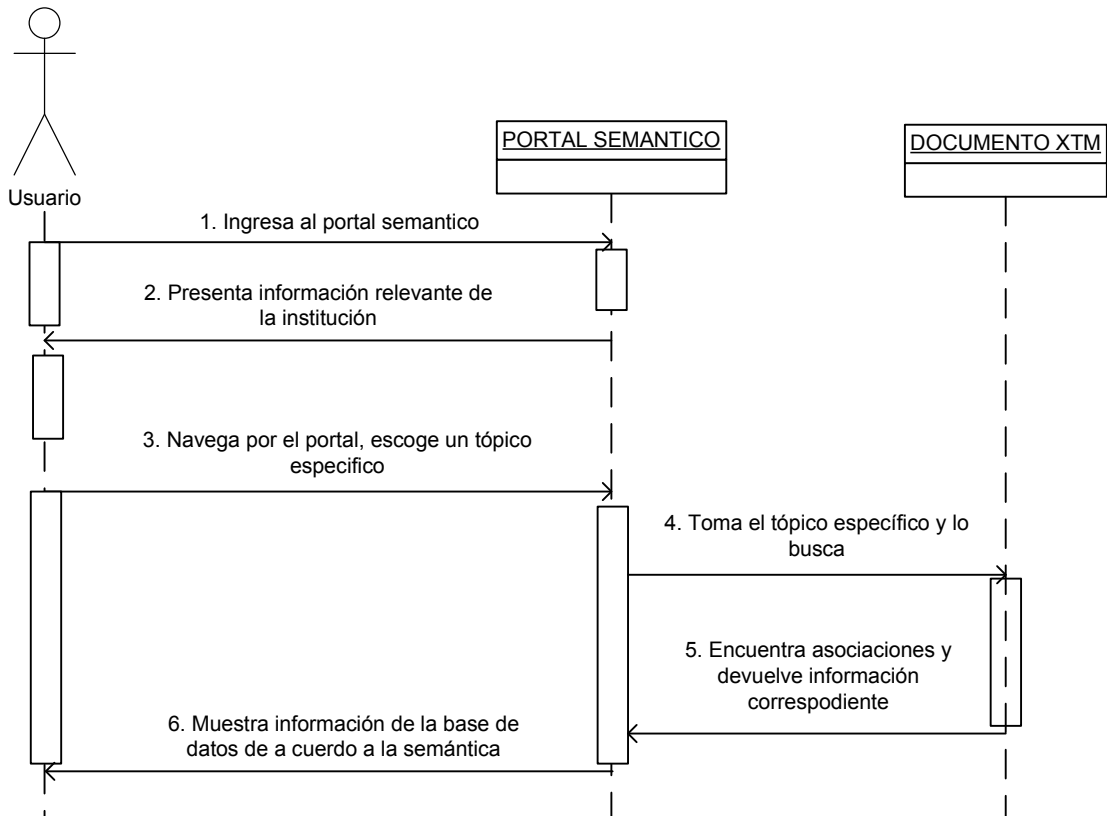


Figura IV.13. Representación gráfica de la historia de usuario2

MÓDULO DE FUNCIONALIDAD DEL USUARIO

Historia de usuario 3

Descripción de las Interacciones:

ACTORES	SISTEMA
1. El Usuario inicia la iteración ingresando a la página de la institución	2. Presenta la página principal
	3. Muestra información general de la institución
4. El usuario escoge las diferentes opciones	5. Muestra la información solicitada y organizada de forma semántica

Tabla IV.31. Representación de la historia de usuario3

REPRESENTACION GRAFICA.

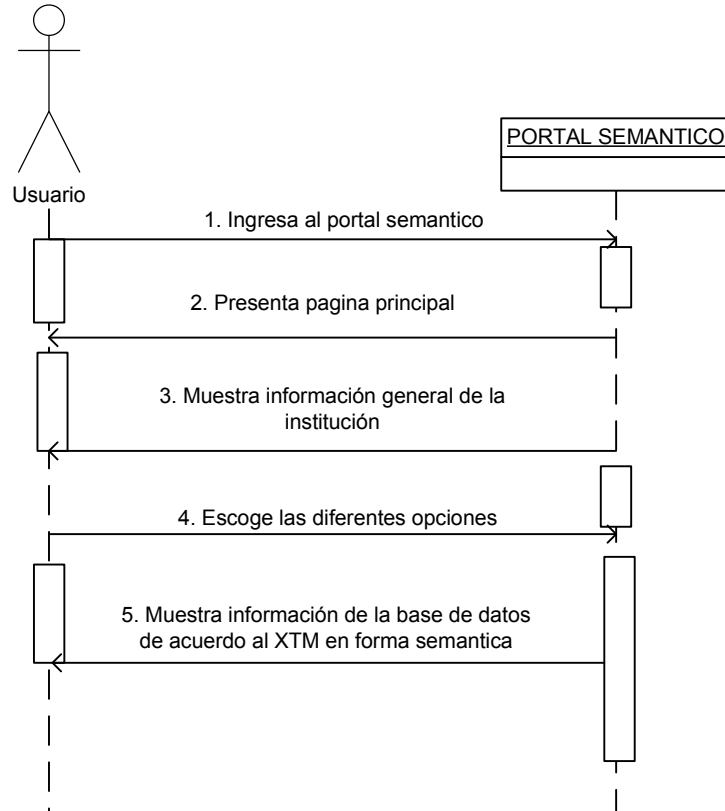


Figura IV.14. Representación gráfica de la historia de usuario3

4.2.5. DIAGRAMA DE CLASES (MODELO CONCEPTUAL)

4.2.5.1 Identificar Clases y Objetos del Ámbito de Análisis

SISTEMA PROPUESTO

Personas

Usuarios

Documentos

Paginas

4.2.5.2 Identificación de Relaciones

Los usuarios que ingresan en las diferentes páginas de la institución recibirán información específica con respecto a un tópico específico, la información enviada desde el documento XTM, será organizada y abstraída de la base de datos de la institución.

4.2.5.3. Diagrama Del Modelo Conceptual

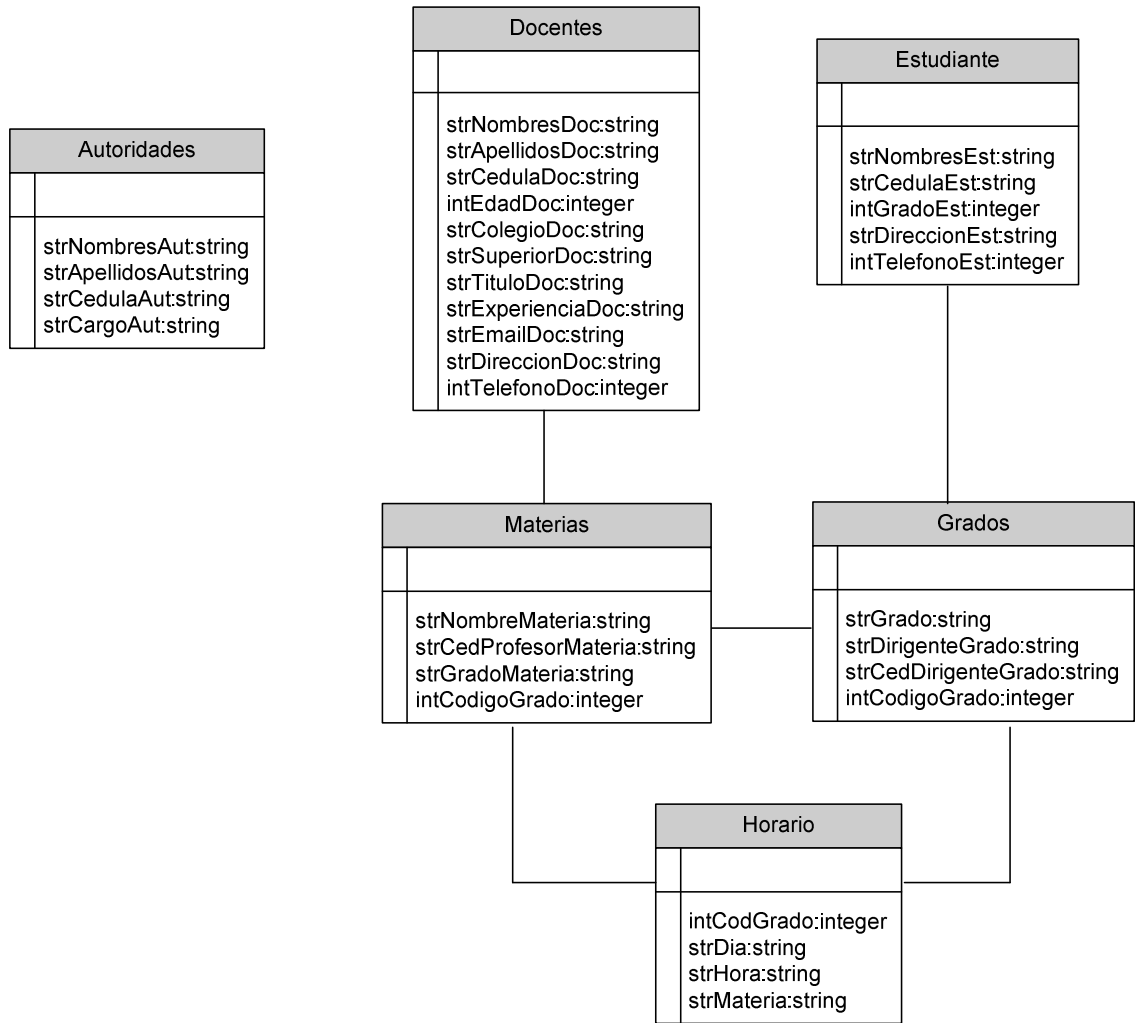


Figura IV.15. Diagrama del modelo conceptual

4.2.6. DIAGRAMAS DE CALLE

4.2.6.1. MÓDULO DE INTEGRACIÓN DE BASE DE DATOS

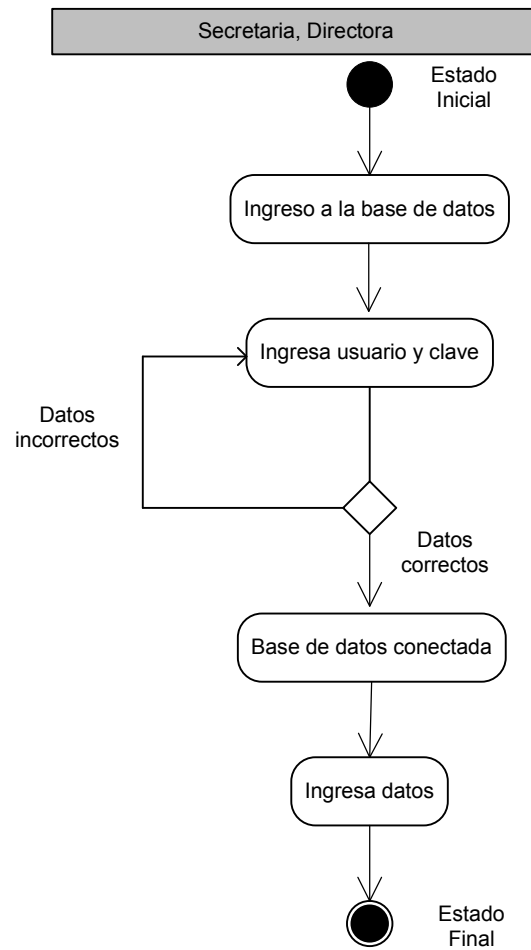


Figura IV.16. Diagrama de calle del módulo 1

4.2.6.2. MÓDULO DE BÚSQUEDA SEMÁNTICA DE INFORMACIÓN

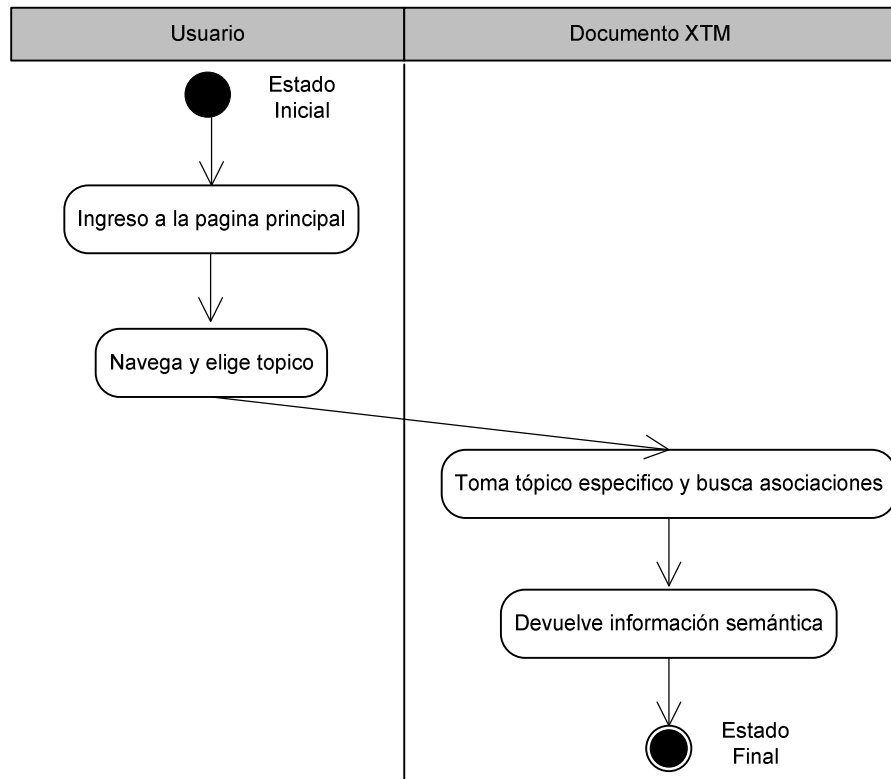


Figura IV.17. Diagrama de calle del modulo2

4.2.6.3. MÓDULO DE FUNCIONALIDAD DEL USUARIO

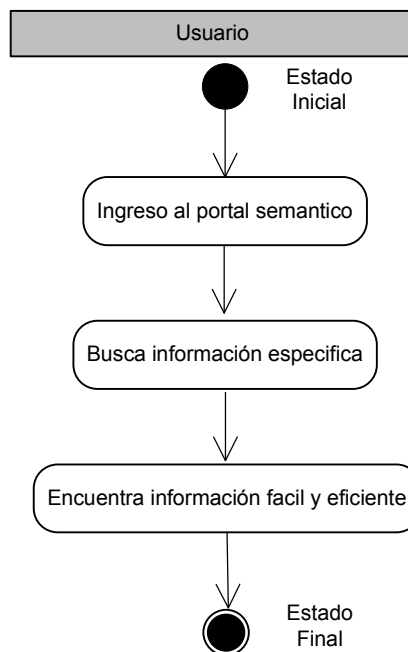


Figura IV.18. Diagrama de calle del modulo3

4.3. FASE DE CODIFICACIÓN

4.3.1. CAPA DE ACCESO A DATOS

La capa de acceso a datos es la que permitirá manejar los datos de la aplicación a través de la interfaz gráfica y la conexión para acceder a los datos que contiene la base de datos.

Esta capa contiene:

Conexión a la base de datos.

Diseño del Portal Semántico.

4.3.1.1. ITERACIONES

PRIMERA ITERACIÓN

ESTRUCTURA DE DIRECTORIOS Y REPOSITORIO DE CÓDIGO FUENTE

Para comenzar el desarrollo es necesario configurar una estructura de directorios adecuada que facilite las tareas. Con el objeto de fomentar la reusabilidad el código se separará en módulos basándose en funcionalidad, y dentro de cada módulo se creará un subproyecto para las capas de aplicación, capa de acceso a datos y capas de reglas de negocio.

FUNCIONALIDAD COMÚN

Dependiendo del modulo se irán añadiendo las funcionalidades que se prevé serán utilizadas más de una vez para facilitar su reusabilidad. Es necesario extraer ciertas tareas repetitivas que serán comunes a gran parte de las clases C#. Se diferencian dos objetivos claros: facilitar el depurado e implementación de métodos definidos en el contrato C#.

Configuración para la conexión a bases de datos.

Esta función se manda a llamar dentro de cada procedimiento en el que se necesite abrir la conexión a la base de datos, este código se coloca en el WebConfig del Sitio Web.

<appSettings>

```
<add key="Cadena Conexion" value="Data Source=Liceo;User
ID=monica;Unicode=True; password=123456" />
</appSettings>
```

SEGUNDA ITERACIÓN

En esta segunda iteración se gestionará la integración de la base de datos, e ingreso de datos a la misma.

Ingreso a la base de datos 2

ESTIMACIÓN INICIAL 2

REAL 2

DIAGRAMA UML

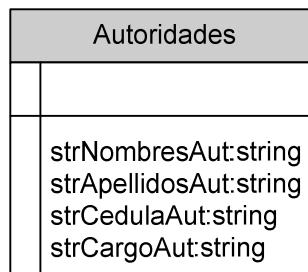


Figura IV.19. Diagrama UML de la segunda iteración

INTEGRACIÓN DE BASE DE DATOS

MODELO

Proporcionará procesos con las que la información de usuario puede estar en memoria, útil en entornos de prueba, en una base de datos accesible vía ODBC. Se realizara el ingreso datos de la institución, para lo cual primero se deben identificar con un usuario y clave que validara el propio manejador de base de datos.

CONFIGURACIÓN DE INTEGRACIÓN DE BASE DE DATOS

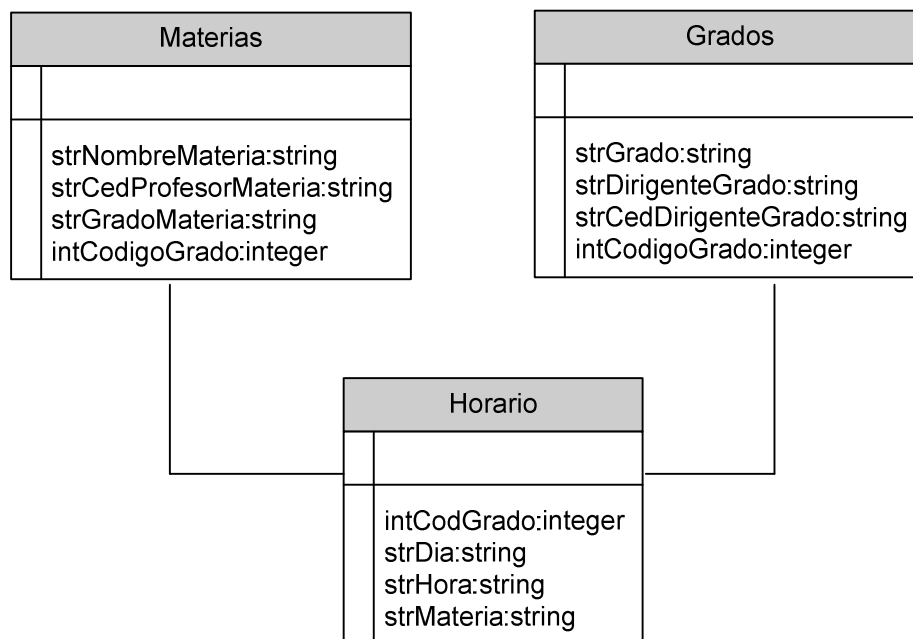
Vamos a utilizar SQL Server 2005 para la creación de la base de datos, creando nuestras tablas con sus respectivos campos y procedimientos almacenados para el ingreso y abstracción de los datos.

TERCERA ITERACIÓN

En esta función se gestionara todos los Menús, GridView que tiene la interfaz y por los cuales se mostrará la información que contiene la aplicación.

Menús y GridView	3
ESTIMACIÓN INICIAL	3
REAL	3

DIAGRAMA UML



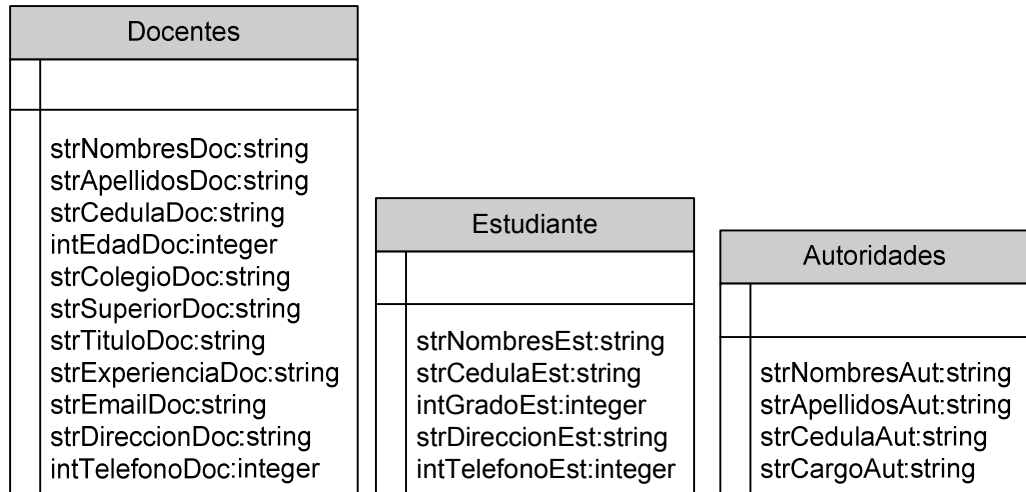


Figura IV.20. Diagrama UML de la tercera iteración

APLICACIÓN WEB GENERAL

Aplicación LiceoStephenHawking

La aplicación web estará formada por las páginas en .NET, documentos XML, los mensajes de la aplicación y los ficheros de configuración.

Interfaz web

Se va a realizar:

Menús y GridViews de la aplicación

Páginas ASP necesarias para proporcionar la funcionalidad de mostrar información semántica.

Páginas de soporte: cabecera, pie de página, menús,...

Mensajes

Mensajes comunes para todo el sistema.

Configuración

Descriptor de aplicaciones web

Acciones y formularios

Validación de los formularios

Menús.

GridViews.

CUARTA ITERACIÓN

BÚSQUEDA SEMÁNTICA DE INFORMACIÓN

En esta cuarta iteración se implementará la funcionalidad relativa a la búsqueda semántica.

Control de datos de tipo tópico específico en XTM	2
Información semántica	2
Información general	1
ESTIMACIÓN INICIAL	5
REAL	5

DIAGRAMA UML

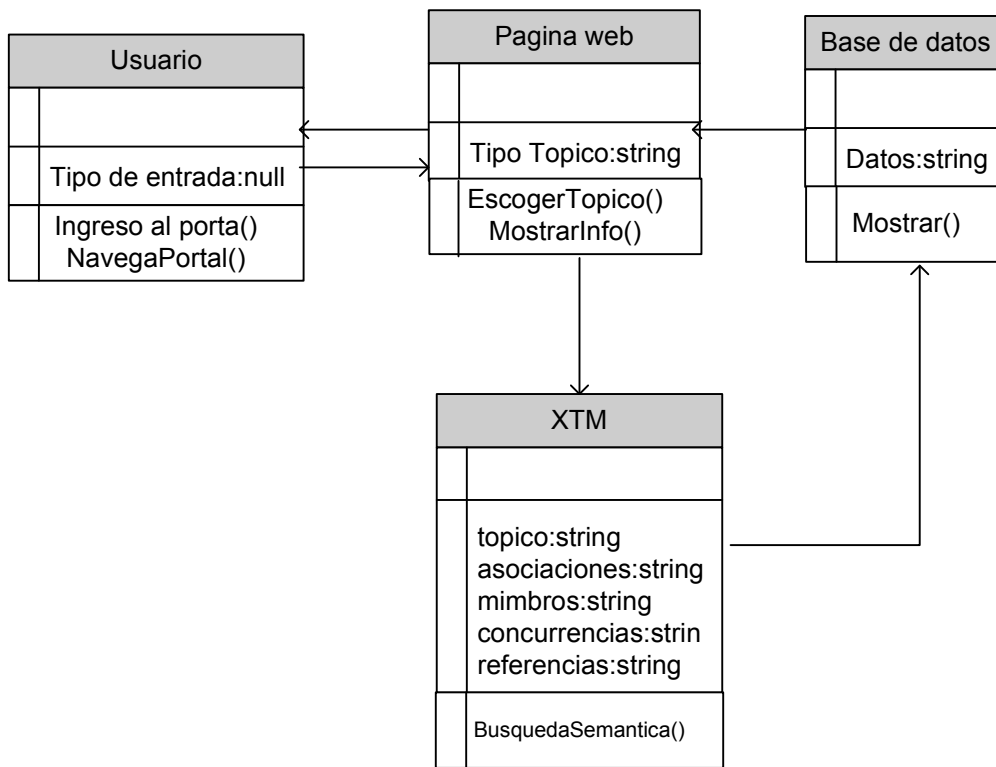


Figura IV.21. Diagrama UML de la cuarta iteración

APLICACIÓN WEB

La aplicación web estará formada por las páginas en .NET, documentos XML, los mensajes de la aplicación y los ficheros de configuración.

Interfaz web

Se va a realizar:

Menús de la aplicación

Búsquedas semánticas en archivos XML

Páginas ASP necesarias para proporcionar la funcionalidad de mostrar información.

Páginas de soporte: cabecera, pie de página, menús,...

Mensajes

Mensajes comunes para todo el sistema.

Configuración

Descriptor de aplicaciones web

Acciones y formularios

Validación de los formularios

Menús.

Búsqueda semántica.

QUINTA ITERACION

FUNCIONALIDAD DEL USUARIO

Esta parte contiene el módulo de funcionalidad de las actividades que puede realizar el usuario donde se le presenta una interfaz amigable donde el puede acceder a la información de la institución y encontrar la información requerida sin olvidar el tópico principal, navegando en forma semántica.

Verificar acceso al portal	1
Navegar semánticamente	1
Encontrar información general	1
ESTIMACIÓN INICIAL	3
REAL	3

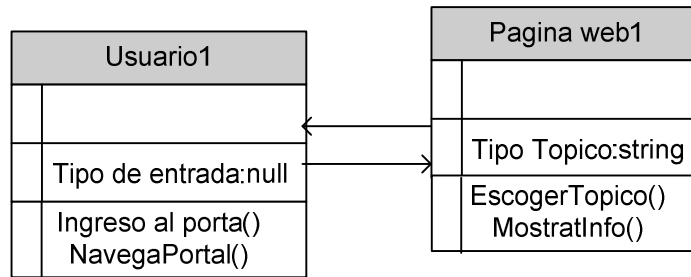
DIAGRAMA UML

Figura IV.22. Diagrama UML de la quinta iteración

APLICACIÓN WEB.

La aplicación web estará formada por las páginas en .NET, documentos XML, los mensajes de la aplicación y los ficheros de configuración.

Interfaz web

Se va a realizar:

Menús de la aplicación

GridViews para mostrar información

Páginas ASP necesarias para proporcionar la funcionalidad al usuario.

Páginas de soporte: cabecera, pie de página, menús,...

Mensajes

Mensajes comunes para todo el sistema.

Configuración

Descriptor de aplicaciones web

Acciones y formularios

Validación de los formularios

Menús.

GridViews.

4.3.2. CAPA DE APLICACIÓN E INTERFAZ**CODIGO PARA LA BÚSQUEDA DE UN TOPICO ESPECÍFICO EN EL DOCUMENTO XTM Y MUESTRA DE SEMANTICA EN LISTBOX**

//cuando la página se carga

```
protected void Page_Load(object sender, EventArgs e)
```



```

{
    if (!Page.IsPostBack)
    {
        lbxSemantica.Items.Clear();

        Session["Topico"] = "instituto";
        string top = Session["Topico"].ToString();
        clTopicMap objTop = new clTopicMap();
        int numeroAsoc;
        int numeroMiembro;
        //buscar en XTM
        string busqSem = BusquedaSemantica(top, out numeroAsoc, out
numeroMiembro);
        string busqSemantica = "";
        if ((numeroAsoc != -1) && (numeroMiembro != -1))
        {
            string unionSemantic = "";
            string Miembro = UnionSemantica(numeroAsoc);
            int numerMiembroAsoc =
objTop.ContarMiembros(numeroAsoc);
            for (int i = 0; i < numerMiembroAsoc; i++)
            {
                if (busqSem != "")
                {
                    busqSemantica =
objTop.ReferenciaRolMiembroAsociacion(numeroAsoc, i);
                    unionSemantic = top + " " + Miembro + " " +
busqSemantica;
                    lbxSemantica.Visible = true;
                }
                lbxSemantica.Items.Add(unionSemantic);
            }
        }
    }
}

```

```

    }
    else
    {
        Page.Response.Redirect("/paginas/instituto/wfInstituto.aspx");
    }
}
}

```

//busqueda semantica

```

public string BusquedaSemantica(string topico, out int numAso, out int
numMiemb)
{
    string buscar = "";
    clTopicMap objTopicMap = new clTopicMap();
    int contarTopic = objTopicMap.ContarTopicos();
    int contarAsociaciones = objTopicMap.contarAsociaciones();
    int Aso = -1;
    int miemb = -1;
    int numTopic;
    string TopicoEncontrado = "";
    string InstanciaTopicRef = "";
    string Asociado = "";
    if (topico != "")
    {
        //busca el topico en el documento XTM
        for (int i = 0; i < contarTopic; i++)
        {
            TopicoEncontrado = objTopicMap.EncontrarTopico(topico, out
numTopic);
            if (TopicoEncontrado == topico)

```

```

    {
        //busca la referencia del topico especifico
        InstanciaTopicRef =
objTopicMap.EncontrarInstanciaTopico(numTopic);
        if (InstanciaTopicRef != "")
        {
            //busca a la asociacion
            for (int j = 0; j < contarAsociaciones; j++)
            {
                int contarMiembros = objTopicMap.ContarMiembros(j);
                if (contarMiembros != 0)
                {
                    for (int k = 0; k < contarMiembros; k++)
                    {
                        Asociado =
objTopicMap.ReferenciaRolMiembroAsociacion(j, k);
                        if (Asociado == InstanciaTopicRef)
                        {
                            buscar = Asociado;
                            Aso = j;
                            miemb = k;
                            break;
                        }
                    }
                    //break;
                }
            } //break;
        }
        break;
    }
}
}

```

```

        numAso = Aso;
        numMiemb = miemb;
        return buscar;
    }

```

4.3.3.CAPA DE REGLAS DE NEGOCIO

La capa de negocios contiene todas las clases utilizadas en nuestra aplicación. Y son siete:

- clAutoridad
- clDocente
- clEstudiante
- clGrado
- clHorario
- clMateria
- clTopicMap

A continuación se muestra la clase clTopicMap:

```

public class clTopicMap
{
    XmlDocument doc = new XmlDocument();
    //encuentra el Id de UN topico especifico
    public string EncontrarTopico(string Topico, out int numTopico)
    {
        string topicId = "";
        int num=0;
        doc.Load("C:/Documents and Settings/Pame/Mis documentos/Visual
Studio
2008/Projects/LiceoStephenHawking/LiceoStephenHawking/App_Data/TopicM
ap.xml");
        XmlNodeList topicMap = doc.GetElementsByTagName("topicMap");
        XmlNodeList topico =
((XmlElement)topicMap[0]).GetElementsByTagName("topic");
        foreach (XmlElement nodo in topico)
        {

```

```

        topicId = nodo.GetAttribute("id");
        num++;
        if(topicId==Topico)
        {
            break;
        }
    }
    numTopico = num-1;
    return topicId; }

//encuentra la instancia-topicRef de un topico especifico
public string EncontrarInstanciaTopico(int numeroTopico)
{
    string topicRef = "";
    string concatenar = "";
    doc.Load("C:/Documents and Settings/Pame/Mis documentos/Visual
Studio
2008/Projects/LiceoStephenHawking/LiceoStephenHawking/App_Data/TopicM
ap.xml");
    XmlNodeList topicos = doc.GetElementsByTagName("topicMap");
    XmlNodeList lista =
((XmlElement)topicos[0]).GetElementsByTagName("topic");
    XmlNodeList instancia =
((XmlElement)lista[numeroTopico]).GetElementsByTagName("instanceOf");
    if (instancia[0] != null)
    {
        XmlNodeList referencia =
((XmlElement)instancia[0]).GetElementsByTagName("topicRef");
        foreach (XmlElement nodo in referencia)
        {
            topicRef = nodo.GetAttribute("xlink:href");
        }
        int longitud = topicRef.ToString().Length;

```

```

    for (int i = 1; i < longitud; i++)
    {
        concatenar = concatenar + topicRef[i].ToString();
    }
    if (concatenar == "")
    {
        concatenar = "no tiene instancia";
    }
}
return concatenar; }

//encuentra la instancia-topicRef de una asociacion especifica
public string EncontrarInstanciaAsociacion(int numeroAsociacion)
{
    string topicRef = "";
    doc.Load("C:/Documents and Settings/Pame/Mis documentos/Visual
Studio
2008/Projects/LiceoStephenHawking/LiceoStephenHawking/App_Data/TopicM
ap.xml");
    XmlNodeList topicos = doc.GetElementsByTagName("topicMap");
    XmlNodeList lista =
((XmlElement)topicos[0]).GetElementsByTagName("association");
    XmlNodeList instancia =
((XmlElement)lista[numeroAsociacion]).GetElementsByTagName("instanceOf"
);
    XmlNodeList referencia =
((XmlElement)instancia[0]).GetElementsByTagName("topicRef");

    //XmlNodeList topicMap = doc.GetElementsByTagName("topicMap");
    //XmlNodeList asociacion =
((XmlElement)topicMap[0]).GetElementsByTagName("association");

```

```

        //XmlNodeList instancia =
((XmlElement)asociacion[numeroAsociacion]).GetElementsByTagName("insta
nceof");
        //XmlNodeList referenciaAsociacion =
((XmlElement)instancia[0]).GetElementsByTagName("topicRef");
        foreach (XmlElement nodo in referencia)
        {
            topicRef = nodo.GetAttribute("xlink:href");
        }
        int longitud = topicRef.ToString().Length;
        string concatenar = topicRef[1].ToString();
        for (int i = 2; i < longitud; i++)
        {
            concatenar = concatenar + topicRef[i].ToString();
        }
        if (concatenar == "")
        {
            concatenar = "no tiene instancia";
        }
        return concatenar;
    }
    //encuentra la referencia a la que hace el rol de un miembro
    public string ReferenciaRolMiembroAsociacion(int numeroAsociacion, int
numeroMiembro)
    {
        string rolMiembro = "";
        doc.Load("C:/Documents and Settings/Pame/Mis documentos/Visual
Studio
2008/Projects/LiceoStephenHawking/LiceoStephenHawking/App_Data/TopicM
ap.xml");
        XmlNodeList topicMap = doc.GetElementsByTagName("topicMap");

```

```

        XmlNodeList asociacion =
((XmlElement)topicMap[0]).GetElementsByTagName("association");
        XmlNodeList miembro =
((XmlElement)asociacion[numeroAsociacion]).GetElementsByTagName("mem
ber");
        XmlNodeList referenciaMiembro =
((XmlElement)miembro[numeroMiembro]).GetElementsByTagName("topicRef
");

        foreach (XmlElement nodo in referenciaMiembro)
        {
            rolMiembro = nodo.GetAttribute("xlink:href"); }
        int longitud = rolMiembro.ToString().Length;
        string concatenar = "";
        for (int i = 1; i < longitud; i++)
        { concatenar = concatenar + rolMiembro[i].ToString(); }
        if (concatenar == "")
        { concatenar = "no tiene instancia"; }
        return concatenar; }
    }

```

DISEÑO RELACIONAL DE LA BASE DE DATOS

4.3.4. MODELADO FISICO

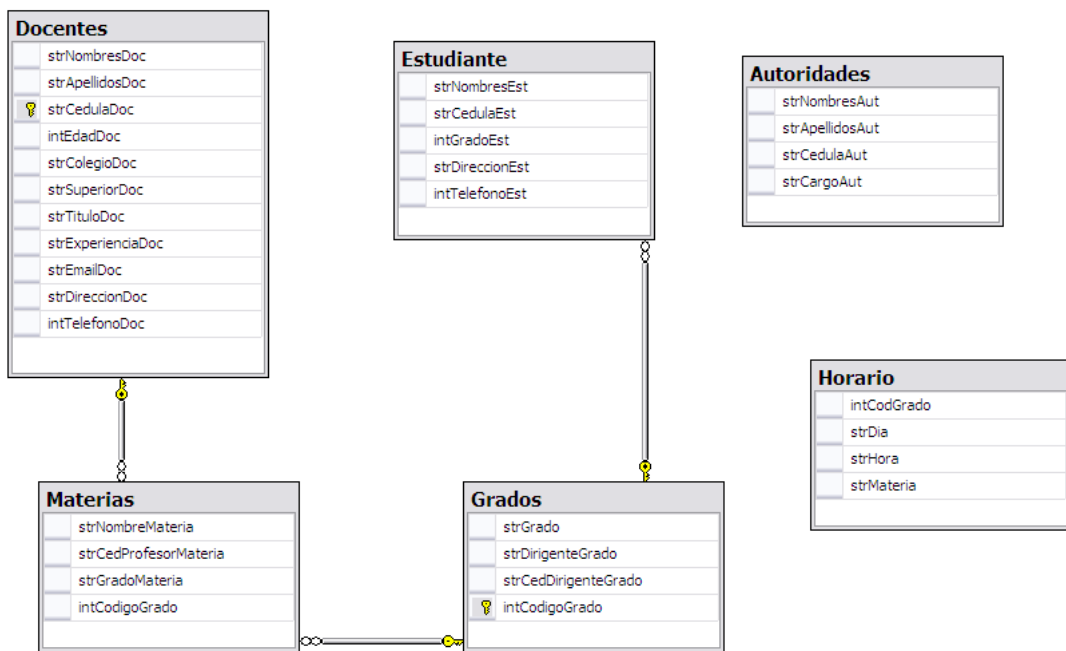


Figura IV.23. Diseño relacional de la Base de Datos

4.3.5. DIAGRAMA DE COMPONENTES

IDENTIFICAR COMPONENTES Y SUS RELACIONES

COMPONENTES

Usuario, Directora, Secretaria.

RELACIONES

La Directora, Secretaria ingresarán a la base de datos “Liceo” para ingresar los datos de los estudiantes, docentes, materias, grados y horarios.

El Portal Semántico mostrará a los usuarios la información de la base de datos según corresponda en el XTM.

CLASIFICAR COMPONENTES

A continuación se clasificarán todos los componentes en tres capas importantes:

Capa Browser

Capa Servidor de Aplicación web

Capa Servidor de Datos

Capa Browser

Usuarios

Capa Servidor de Aplicación web

Usuarios

Interfaz de Información General del Liceo Stephen Hawking.

Interfaz de Información de Autoridades del Liceo Stephen Hawking.

Interfaz de Información de Docentes del Liceo Stephen Hawking.

Interfaz de Información de Pensum del Liceo Stephen Hawking.

Interfaz de Información de Estudiantes del Liceo Stephen Hawking.

Capa del servidor de Datos

Acceso a la Base de Datos

Control y Análisis

Rutinas de Conexión

DIAGRAMAS DE COMPONENTES

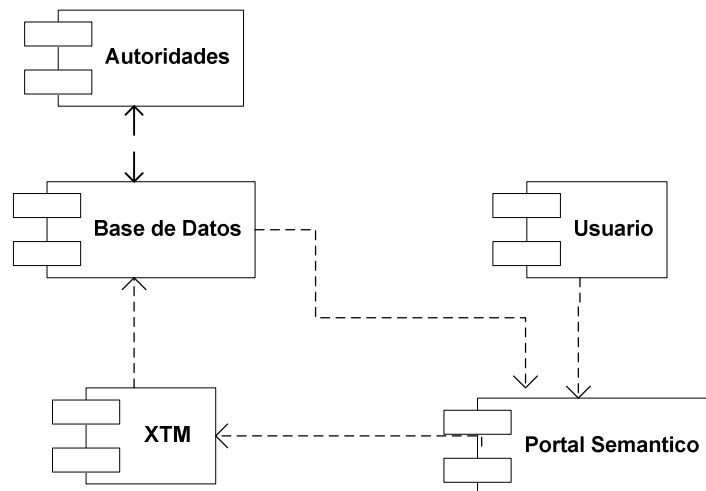


Figura IV.24. Diagrama de Componentes

ANALIZAR LA ESTRUCTURA FÍSICA A TRAVÉS DE LOS NODOS

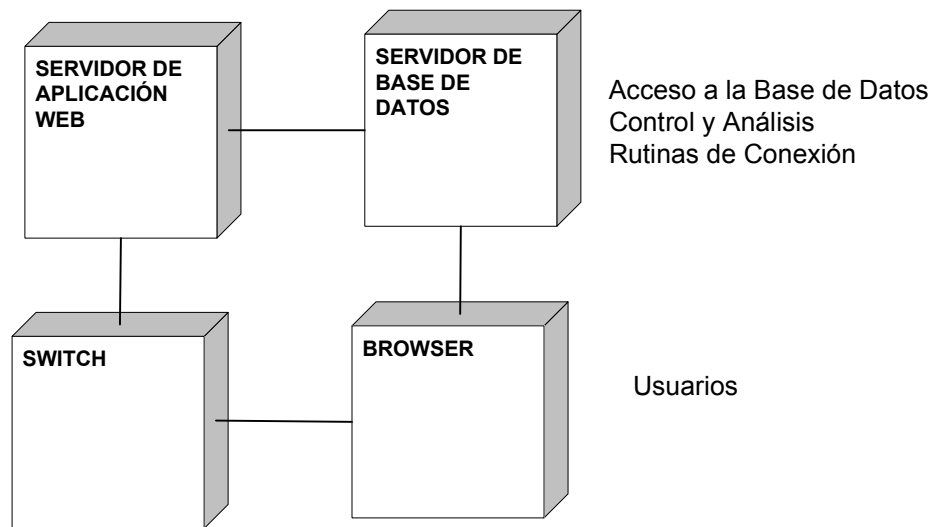


Figura IV.25. Estructura física de los nodos

4.4. FASE DE PRUEBAS

Uno de los pilares de la metodología X.P es el uso de test para comprobar el funcionamiento de los códigos que vayamos implementando, en esta fase se deben crear las aplicaciones que realizarán los test con un entorno de desarrollo específico para test. Hay que someter a tests las distintas clases del sistema omitiendo los métodos más triviales. Se deben crear los test que pasarán los códigos antes de implementarlos. Se requiere de pruebas adicionales y revisiones de rendimiento antes de que el sistema sea trasladado al entorno del cliente. Al mismo tiempo, se deben tomar decisiones sobre la inclusión de nuevas características a la versión actual, debido a cambios durante esta fase.

4.4.1. PRESENTACIÓN DE LOS TEST FINALIZADOS

De acuerdo a las gestiones mencionadas anteriormente en la fase de análisis presentamos los siguientes test con pruebas satisfactorias para el cliente:

GESTIÓN DE INFORMACIÓN GENERAL

Para esta gestión se realizó las siguientes pruebas de acuerdo a los requerimientos funcionales presentados.

	Caso	Resultado	Quien	Perfil	Fecha prueba	Corrección	Fecha corrección	Quien ejecutó la corrección
RF1	La base de datos deberá presentar un recuadro para logueo pidiendo usuario y contraseña.	Exitoso	Pamela Fuertes	Developer	09/09 /2009			
RF2	La base de datos deberá validar los datos ingresados y abrir la base de datos en caso de éxito, o no acceder en caso de fracaso.	Exitoso	Pamela Fuertes	Developer	09/09 /2009			
RF3	En caso de éxito los datos deberán llenarse y validarse antes de la presentación del portal	Exitoso	Pamela Fuertes	Developer	10/09 /2009			

	a los usuarios.							
--	-----------------	--	--	--	--	--	--	--

Tabla IV.32. Prueba de la gestión de información general

	Caso	Resultado	Quien	Perfil	Fecha prueba	Corrección	Fecha corrección	Quien ejecutó la corrección
RF4	El portal presenta información relevante de la institución en la parte central del portal y al lado izquierdo la navegación de acuerdo al tema a escoger.	Exitoso	Pamela Fuertes	Developer	21/09 /2009			
RF5	El portal permitirá navegar al usuario de acuerdo a la semántica mostrada, dejándolo escoger un	Fallido	Pamela Fuertes	Developer	28/09 /2009	Se habilitará los botones para la navegación	28/09 /2009	Pamela Fuertes
		Exitoso	Pamela Fuertes	Developer	28/09 /2009			

	tema específico.							
RF6	El portal guarda el tema específico y lo lleva buscar dentro del documento XTM.	Fallido	Pamela Fuertes	Developer	29/09 /2009	Se crearán los procedimientos necesarios para la búsqueda en documento XTM	07/10 /2009	Pamela Fuertes
		Exitoso	Pamela Fuertes	Developer	07/10 /2009			
RF7	El portal encuentra la o las asociaciones para el tema escogido, y devuelve la información específica de acuerdo a la semántica presentada.	Exitoso	Pamela Fuertes	Developer	09/10 /2009			
RF8	El portal busca la información en la base de datos de acuerdo a la información	Exitoso	Pamela Fuertes	Developer	12/10 /2009			

n solicitada y la muestra al usuario.								
--	--	--	--	--	--	--	--	--

Tabla IV.33. Prueba de la gestión de información semántica

Para las pruebas se tomaron en cuenta los requerimientos funcionales de acuerdo a las gestiones; en la gestión de información general se obtuvieron resultados exitoso, y en la gestión de información semántica lo requerimientos RF5 y RF6 hubo algunos fallos mismo que se resolvieron y se volvieron a ejecutar con éxito.

ANÁLISIS DE RESULTADOS DE LA HIPÓTESIS

Este análisis es con el fin de realizar la validación de la hipótesis de el “Desarrollo de un portal semántico bajo el paradigma XML Topic Maps (XTM): Caso práctico Liceo Stephen Hawking.”, se determina que la estadística inferencial y específicamente el Chi cuadrado es el método más adecuado, puesto que se aplica a pequeñas poblaciones y se aplica debido a que no existe un sistema anterior.

ESTADISTICA INFERENCIAL

La inferencia estadística o estadística inferencial es una parte de la Estadística que comprende los métodos y procedimientos para deducir propiedades (hacer inferencias) de una población, a partir de una pequeña parte de la misma (muestra).

La bondad de estas deducciones se mide en términos probabilísticos, es decir, toda inferencia se acompaña de su probabilidad de acierto.

SELECCIÓN Y DETERMINACIÓN DE LA MUESTRA

Se realizó una encuesta para poder obtener información necesaria para la comprobación de la hipótesis, misma que se hizo a un grupo de personas que entendían de portales web comunes, y se les hizo las preguntas respectivas, dichas preguntas eran claras y rápidas lo que evitó malestar y confusión a los encuestados.

La encuesta se encuentra en el anexo 1.

OBTENCIÓN DE LOS DATOS

La encuesta se la realizó a veinte personas que habían utilizado en primera instancia portales web comunes; después de haber utilizado en portal del Liceo de Talentos Stephen Hawking se les hizo ocho preguntas fundamentales para la demostración de la hipótesis, cuatro con respecto a la navegabilidad y cuatro con respecto a la eficiencia.

- La creación de un ambiente amigable como un portal semántico ayuda a desplazarse de mejor manera en la búsqueda de información.
- La creación de un ambiente amigable como un portal semántico NO le ayuda a desplazarse de mejor manera en la búsqueda de información.
- Sin la creación de un ambiente amigable como un portal semántico usted piensa que le ayuda a desplazarse de mejor manera en la búsqueda de información.
- Sin la creación de un ambiente amigable como un portal semántico usted piensa que NO le ayuda a desplazarse de mejor manera en la búsqueda de información.
- La creación de un ambiente amigable como un portal semántico le muestra los resultados deseados mediante los recursos disponibles.
- La creación de un ambiente amigable como un portal semántico NO le muestra los resultados deseados mediante los recursos disponibles.
- Si la creación de un ambiente amigable como un portal semántico le muestra los resultados deseados mediante los recursos disponibles.
- Sin la creación de un ambiente amigable como un portal semántico NO le muestra los resultados deseados mediante los recursos disponibles.

La calificación de las preguntas a las respuestas está entre uno y cinco (1-5) tomando a uno (1) como la calificación menor y cinco (5) como la calificación mayor.

Sin embargo el encuestado no está en la obligación de responder a todas las preguntas por lo que hemos tomado el valor de 0 (cero) a las preguntas no respondidas.

CLASIFICACIÓN Y ORGANIZACIÓN DE LOS DATOS

Las respuestas emitidas por los encuestados son las siguientes:

Encuestados	Enunciados							
	Preg1	Preg2	Preg3	Preg4	Preg5	Preg6	Preg7	Preg8
1	3	2	0	4	4	1	2	1
2	4	1	2	3	5	2	0	2
3	5	1	0	3	5	2	1	2
4	4	0	2	3	3	1	2	3
5	4	2	0	4	3	1	1	3
6	4	0	2	3	4	2	1	2
7	3	1	0	3	4	2	0	4
8	3	1	0	2	4	1	2	3
9	5	1	1	0	4	1	0	4
10	5	2	2	3	5	1	0	5
11	4	0	1	0	3	2	0	2
12	4	2	1	3	3	1	1	2
13	2	2	0	3	4	1	0	4
14	3	1	2	0	1	1	1	3
15	3	2	0	2	5	2	0	4
16	2	1	1	2	2	0	0	2
17	4	2	1	3	1	1	2	3
18	5	0	1	3	3	1	2	5
19	4	2	1	3	5	1	1	2
20	5	0	2	3	4	0	0	4
TOTAL	76	23	19	50	72	24	16	38

Tabla IV.34. Respuestas a preguntas de la encuesta.

ANÁLISIS DE LOS DATOS

Se muestran los datos de las respuestas de las encuestas en una matriz dada por filas (i) y columnas (j), en una matriz para mayor facilidad al momento de aplicar Chi-cuadrado.

	ayuda a desplazarse de mejor manera en la búsqueda de información	No ayuda a desplazarse de mejor manera en la búsqueda de información	TOTAL
La creación de un ambiente amigable como un portal semántico	76	23	99
Sin la creación de un ambiente amigable como un portal semántico	19	50	69
TOTAL	95	73	168

Tabla IV.35. Valores abstraídos en base a las cuatro primeras preguntas

	Le muestra los resultados deseados mediante los recursos disponibles	NO le muestra los resultados deseados mediante los recursos disponibles	TOTAL
La creación de un ambiente amigable como un portal semántico	72	24	96
Si la creación de un ambiente amigable como un portal semántico	26	60	86
TOTAL	98	84	182

Tabla IV.36. Valores abstraídos en base a las cuatro últimas preguntas

Ahora se va a determinar los valores esperados, a partir de los datos abstraídos, para esto se debe multiplicar los respectivos marginales y dividir por el gran total.

Para las cuatro primeras preguntas de la *Tabla 4.35*:

$$V.E = (\sum \text{fila (i)} * \sum \text{columna (j)}) / \text{total}$$

$$V.E_{11} = (99*95)/168 = 55,98$$

$$V.E_{12} = (99*73)/168 = 43,01$$

$$V.E_{21} = (69*95)/168 = 39,01$$

$$V.E_{22} = (69*73)/168 = 29,98$$

Para las cuatro últimas preguntas de la *Tabla 4.36*:

$$V.E = (\sum \text{fila (i)} * \sum \text{columna (j)}) / \text{total}$$

$$V.E_{11} = (96*98)/182 = 51,69$$

$$V.E_{12} = (96*84)/182 = 44,30$$

$$V.E_{21} = (86*98)/182 = 46,30$$

$$V.E_{22} = (86*84)/182 = 39,69$$

VALORES ESPERADOS

	ayuda a desplazarse de mejor manera en la búsqueda de información	No ayuda a desplazarse de mejor manera en la búsqueda de información	TOTAL
La creación de un ambiente amigable como un portal semántico	55,98	43,01	98,99

Sin la creación de un ambiente amigable como un portal semántico	39,01	29,98	68,99
TOTAL	94,99	72,99	167,98

Tabla IV.37. Valores esperados de las cuatro primeras preguntas

	Le muestra los resultados deseados mediante los recursos disponibles	NO le muestra los resultados deseados mediante los recursos disponibles	TOTAL
La creación de un ambiente amigable como un portal semántico	51,69	44,3	95,99
Sin la creación de un ambiente amigable como un portal semántico	46,3	39,69	85,99
TOTAL	97,99	83,99	181,98

Tabla IV.38. Valores esperados de las cuatro últimas preguntas

A través de Chi-cuadrado se probará de forma afirmativa o negativa que la distribución de las frecuencias abstraídas difiere significativamente en relación a la distribución de las frecuencias que deberíamos esperar.

$$X^2 = (\sum(O-E)^2) / E$$

O = Frecuencia o valores observados

E = Frecuencia o valores Esperados

O	E	O-E	(O-E)²	(O-E)²/E
76	55,98	20,02	400,8	7,15

23	43,01	-20,01	400,4	9,30
19	39,01	-20,01	400,4	10,26
50	29,98	20,02	400,8	13,36
			X² =	40,07

Tabla IV.39 Chi-Cuadrado para las primeras cuatro preguntas

O	E	O-E	(O-E)²	(O-E)²/E
72	51,69	20,31	412,49	7,98
24	44,30	-20,30	412,09	9,30
16	46,30	-30,30	918,09	19,82
60	39,69	20,31	412,49	10,39
			X² =	47,49

Tabla IV.40 Chi-Cuadrado para las últimas cuatro preguntas

Para afirmar o negar la hipótesis se debe comparar el valor obtenido (40,07 y 47,49), con el chi-cuadrado crítico de la tabla de valores crítico, los parámetros que se deben tomar en cuenta son los grados de libertad (GL) y el nivel de significación (P); el primero se define como el (número de columnas-1) X (número de filas-1), en este caso es (2-1) X (2-1) = 1 para ambos casos; el nivel de significación también conocido como nivel de confianza se refiere a la probabilidad de que los resultados observados se deban al azar, este valor es fijado por el investigador usualmente es de 5 % o 10 % (es decir 0,05 o 0,1).

Considerando P = 0,05 y GL=1, se tiene que X² crítico es igual a 3,841 de acuerdo a la Tabla de Valores Críticos de la Distribución Chi-cuadrado estándar. Se observa que X² calculado 40,07 > 3,841 y 47,49 > 3,841.

Se afirma que existe una asociación entre las variables estudiadas por tanto se afirma que la creación de un ambiente amigable como un portal semántico ayuda a desplazarse de mejor manera en la búsqueda de información y además que muestra los resultados deseados mediante los recursos disponibles.

REPRESENTACIÓN GRÁFICA DE LOS RESULTADOS

Se presenta a continuación los puntajes obtenidos en la encuesta, teniendo en cuenta que el primer gráfico es de las primeras cuatro preguntas y el segundo de las cuatro últimas preguntas.

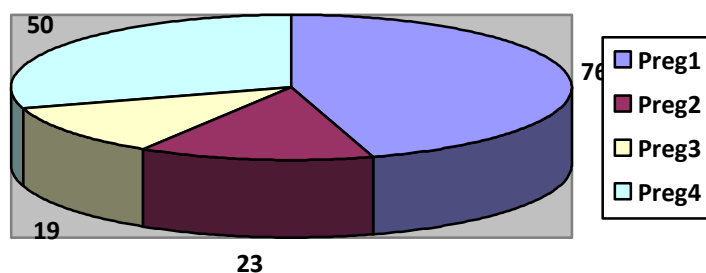


Figura IV.26 Gráfico de resultados de las primeras cuatro preguntas de la encuesta.

Se puede ver que la primera pregunta tiene un total de 76 puntos es decir es la más alta, a continuación la cuarta con 50 puntos y la segunda y tercera por debajo de las mencionadas anteriormente.

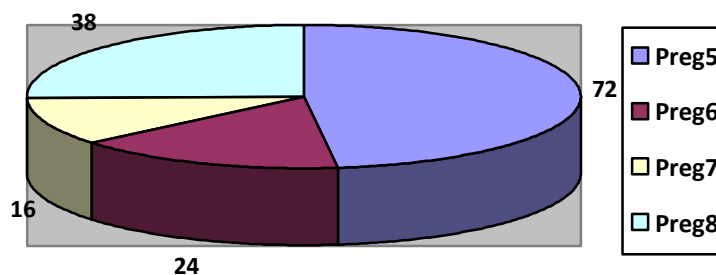


Figura IV.27 Gráfico de resultados de las últimas cuatro preguntas de la encuesta.

Se puede ver que la quinta pregunta tiene un total de 72 puntos es decir es la más alta, a continuación la cuarta con 38 puntos y la segunda y tercera por debajo de las mencionadas anteriormente.

En la encuesta realizada a las veinte personas en algunos enunciados se respondió con cero a lo que se le va a llamar ausencia de votante por lo tanto a continuación se va a detallar el número de votantes y los puntajes obtenidos.

	Enunciados							
	Preg1	Preg2	Preg3	Preg4	Preg5	Preg6	Preg7	Preg8
Votantes	20	15	13	17	20	18	11	20

Tabla IV.41. *Votantes por cada Pregunta*

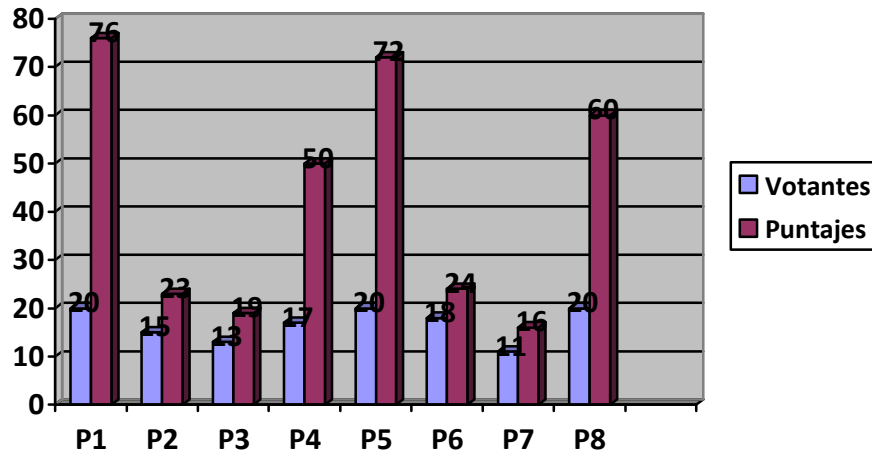


Figura IV.28. *Relación del número de votantes y el puntaje obtenido*

La primera, quinta y octava pregunta son respondidas por todos los votantes, seguida de la cuarta y sexta que la responde casi todos y las demás preguntas son respondidas por menos votantes. A continuación se muestra una gráfica de los votantes y puntajes obtenidos.

Es importante tomar en cuenta que la columna de puntaje que es mucho mayor a la de los votantes es porque los votantes valoraron con puntajes altos a las preguntas, por lo cual las columnas que tienen tamaño similar es porque recibieron el mínimo de calificación y en algunos casos sin calificación.

VALIDACIÓN DE LA HIPÓTESIS

La implementación es el núcleo de una aplicación, utilizamos la ingeniería de software para realizar la investigación para poder llegar a presentar la solución deseada, existen grupos que se encuentran trabajando con la finalidad de disminuir el trabajo que realiza el programador, disminuyendo de esta manera el tiempo

empleado en buscar una respuesta a algo en concreto, sin descuidar características como la calidad, la seguridad, la eficiencia, etc.

Si se puede contar con documentos XTM para generar aplicaciones que nos permiten realizar búsquedas de información navegando de forma ágil y semántica; a la vez obtener la información necesaria de manera eficiente, solamente con los recursos presentados.

CONCLUSION DEL ANÁLISIS DE LA HIPÓTESIS

La encuesta realizada se hizo a personas que habían utilizado Internet en varias ocasiones, para asegurarnos de que la muestra sea real, la encuesta presentada fue rápida y fácil de entender, de estos resultados obtuvimos que el Chi-cuadrado calculado es mayor que el Chi-cuadrado crítico dado por defecto, por lo que podemos confirmar la validez de la hipótesis, y decir que el resultado es efectivo por la naturaleza de los datos obtenidos.

CONCLUSIONES

- Luego de haber realizado un estudio y analizado la representación de la información y el conocimiento con XML Topic Maps (XTM), su funcionamiento y procesamiento, puedo concluir como principal punto que la manera más fácil de navegar, es utilizando XTM, puesto que la semántica de una página web está de manera imperceptible al ojo humano, pero la mente lo reconoce de manera más fácil, ágil y eficiente a la hora de buscar información, sin tener que salirse del tema principal.
- Además es muy fácil su elaboración y se puede procesar bajo cualquier herramienta que soporte documentos XML en cualquier ámbito, por lo cual la elaboración del portal semántico en el Liceo Stephen Hawking fue de gran ayuda.
- Con la implementación del portal semántico en el Instituto a más de darse a conocer en la web de una manera sencilla, y de fácil acceso para cualquier usuario, la inversión económica que será de \$4.659,20; será mínima frente a los beneficios como son:
 - ✓ Aumento en la demanda estudiantil
 - ✓ Mayores ingresos económicos a la Institución; y,
 - ✓ Lo que es más aumento en la demanda social para la formación en cuanto a educación se refiere.

RECOMENDACIONES

Si empieza un proyecto utilizando XTM tenga en cuenta que debe formar un DTD para el documento en primera instancia, para luego proceder con la elección de la herramienta adecuada para el procesamiento de la misma.

Para facilitar el desarrollo del nuevo proyecto utilizando XTM es recomendable utilizar la herramienta que usted domine para mejor manejo de la información y mayor aprovechamiento de su experiencia

Tener en cuenta al crear el documento XTM que se puede utilizar cualquier herramienta de texto el ejemplo más fácil es WordPad o un Bloc de notas simplemente, y debe guardarlo con la extensión .xml, igualmente que este bien elaborado, es decir que siga la nomenclatura correcta, para evitar errores posteriores.

Se recomienda al Liceo Stephen Hawking se establezca un sistema de mantenimiento y actualización de la base de datos y del portal semántico, con el fin de mantener renovada la fuente de información al usuario.

RESUMEN

Se desarrollo un portal web semántico bajo el paradigma XML Topic Maps (XTM) en el Liceo Stephen Hawking, con la finalidad de mejorar su navegabilidad y eficiencia en la entrega de información del instituto.

Se utilizó SQL Server 2005 como base de datos, Visual Studio 2008 como herramienta de desarrollo, lenguaje c#, y bloque de notas; implementando con Programación Orientada a Objetos, se crea un archivo XML con sintaxis XTM que utiliza a su vez una Definición de Tipo de Documento (DTD), que contiene los temas y asociaciones necesarios para estructurar la semántica del portal; se toma un tema específico y se busca en estos documentos con la información devuelta, se envía a la base de datos, y esta devuelve los datos necesarios al usuario.

Aplicándose encuestas se probó la web desarrollada a veinte personas por un lapso de diez minutos, comprobándose que el portal mejoró la navegabilidad en un 80,14% y la eficiencia de la información entregada al usuario en un 94%, logrando que el portal web semántico se encuentre organizado y sea específico.

Se pudo concluir que el desarrollo del portal mejoró la navegabilidad y eficiencia al entregar información, por lo tanto se recomienda la implantación y utilización del portal web semántico cuyas mejoras fueron lo suficientemente probadas

SUMMARY

It was development a semantic web site under the XML Topic Maps paradigm (XTM) at Liceo Stephen Hawking, in order to improve the navigability and efficiency related with the information delivering of Institute.

It was used SQL Server 2005 as database manager, Visual Studio 2008 as development tool, C# language and notepad; implementing O-O programming, it was created a XML file with XTM syntax which used turn a Document Type Definition (DTD), that contains the topics and associations needed to structure the semantic web site; it was taken a specific topic and searched in this documents from the result information, this information is sent to database and the request information is shown to the user.

Applying surveys, twenty people tried the web site for ten minutes, proving that web site improves the navigability in 80.14% and efficiency of delivered information to the user in 94%, making possible semantic web site were organized and specific.

It was possible conclude that the semantic web site development improved the navigability and efficiency for information delivering, for this reason it is recommended the implementation and use of semantic web site which improvements were proved enough.

BIBLIOGRAFÍA

LIBROS

- 1.- DACONTA, M. The Semantic Web: aguide to the future of xml, web services, and knowledge. Indianapolis: Wiley, 2003. pp. 1-14, 27-54,181-231
- 2.- DELANEY, K. A fondo Microsoft SQL Server 2000. Madrid: McGraw Hill, 2001. pp. 86
- 3.- GROFF, J.; WEINBERG, P. Guía LAN Times de SQL. Madrid: McGraw Hill, 1998. pp. 219-235
- 4.- PARK, J., HUNTING, S. XML Topic Map, Creating and using Topic Map for the web. USA: Addison-Wefley, 2002. pp. 17-30, 81-100, 103-147
- 5.- CARABALLO PÉREZ, Y. Los Topic Maps y su Relación con las Redes Sociales. Aeimed. Cuba. 16(4). Abril 2007

DIRECCIONES WEB

- 1.- DÍAZ, E. Web Semánticas.
<http://usuarios.lycos.es/websemanticas/topic%20maps.html>
20090801
- 2.- GRUBER. Introducción a las Ontologías.
http://es.geocities.com/ontologias_tesauros/introduccionontologias.html1993
20090627
- 3.- KNOWLEDGE MASTER CORPORATION Y KNOWLEDGE MASTER. Mapas Conceptuales (redes semánticas) y el Formato XTM: los topic maps o los mapas de contenidos
<http://www.conceptmaps.it/KM-XTM-esp.htm>
20090915

- 4.- LAMARCA, M. XTM Topic Maps.
http://www.hipertexto.info/documentos/topic_maps.htm
20090819
- 5.- MEMBERS OF THE TOPICMAPS.ORG AUTHORIZING GROUP. XML Topic Maps (XTM) 1.0.
<http://www.topicmaps.org/xtm/1.0/index.html>
20090613,
- 6.- MOLINA, B. Maps Topics Introducción al Concepto de los Maps Topics.
http://es.geocities.com/topics_maps/
20090601
- 7.- MOREIRO, J. Mapas Conceptuales, Topic maps y Tesoros.
<http://www.fiv.upv.es/jotri/Ponencias/Mapas.pdf>
20090520
- 8.- MOREIRO, J. De los Tesoros a los Topic Maps: nuevo estándar para la representación y la organización de la información

<Http://www.periodicos.ufsc.br/index.php/eb/article/viewFile/178/5470>
20090628
- 9.- NISHIKAWA, M. Xml, XTM

<http://osdir.com/ml/text.xml.xtm.general/2003-01/msg00125.html>
20090618
- 10.- PINTO MOLINA, M. Ejemplos de Topic Map.

<http://www.mariapinto.es/e-coms/topicmap.htm>
20090522
- 11.- RODRIGUEZ, K., RONDA, R., Web Semántica: un nuevo enfoque para la organización y recuperación de información en el web.

http://bvs.sld.cu/revistas/aci/vol13_6_05/aci030605.htm

20090730

- 12.- TOPICMAPS.ORG. Topic Maps: normativa, aplicaciones y herramientas.
<http://tramullas.com/jaca/gescon/contenidos/textos/Garrido.pdf>

20090710

- 13.- TOPICMAPS.ORG. Accessibility Features in XML Topic Map (XTM) version
1.0 Specification.

<http://www.topicmaps.org/xtm/1.0/xtm1-wai.html>

20090623,

- 14.- TUTORIAL, THE ONTOPIA SCHEMA LANGUAGE.

<http://www.ontopia.net/omnigator/docs/schema/tutorial.html>

20090620

- 15.- VILLALOBOS, J. Proyecto Cumbia.

<http://agamenon.uniandes.edu.co/~jvillalo/doku.php?id=projects>

20090803

- 16.- WORKIN GROUP HOME PAGE. W3C, XHTML2

<http://www.w3.org/MarkUp/>

20090725

ANEXOS

ENCUESTA

De acuerdo a su experiencia en páginas o portales web por favor responda las siguientes preguntas calificándolas entre 1-5 tomando como menor puntaje 1 y mayor puntaje 5, o 0 (cero) en caso de abstenerse:

- La creación de un ambiente amigable como un portal semántico ayuda a desplazarse de mejor manera en la búsqueda de información.

Calificación _____

- La creación de un ambiente amigable como un portal semántico NO le ayuda a desplazarse de mejor manera en la búsqueda de información.

Calificación _____

- Sin la creación de un ambiente amigable como un portal semántico usted piensa que le ayuda a desplazarse de mejor manera en la búsqueda de información.

Calificación _____

- Sin la creación de un ambiente amigable como un portal semántico usted piensa que NO le ayuda a desplazarse de mejor manera en la búsqueda de información.

Calificación _____

- La creación de un ambiente amigable como un portal semántico le muestra los resultados deseados mediante los recursos disponibles.

Calificación _____

- La creación de un ambiente amigable como un portal semántico NO le muestra los resultados deseados mediante los recursos disponibles.

Calificación _____

- Si la creación de un ambiente amigable como un portal semántico le muestra los resultados deseados mediante los recursos disponibles.

Calificación _____

- Sin la creación de un ambiente amigable como un portal semántico NO le muestra los resultados deseados mediante los recursos disponibles.

Calificación _____