

Tesis: "Análisis del algoritmo ActiveRank como método de detección automática de contenido dentro de redes de información" Fernando Luege Mateos



México D.F., Febrero 2010

Índice

	Página
1. Definiciones	1
1.1. ¿Cómo leer esta tesis?	2
1.2. Contexto	2 3 4
1.3. Teoría de Gráficas	4
1.4. Computación y redes	5
1.5. Internet y World Wide Web	6
2. Antecedentes	7
2.1. Conceptos Básicos de Teoría de Gráficas	8
2.1.1. Definición de Teoría de Gráficas	8
2.1.2. Puentes de Köingsberg	9
2.1.3. Conceptos básicos	10
2.2. World Wide Web	13
2.2.1. Internet	13
2.2.2. Inicio y fundamentos de la World Wide Web	13
2.2.3. Hipervínculos	14
2.2.4. Dimensiones de la WWW	16
2.2.5. Análisis de la WWW	17
2.3. Sistemas de Análisis de Redes de Información	20
2.3.1. Estructura general	20
2.3.2. Integración y procesamiento de información	25
2.3.3. Infraestructura de los sistemas de análisis de información	26
2.4. Marco General de la Pornografía en Internet	31
2.4.1. Definición de pornografía	31
2.4.2. Estadísticas del perfil económico y demográfico de la pornografía en Internet	32
2.4.3. Estructura del contenido pornográfico en Internet	36
2.4.4. Legislación y ética de la pornografía en Internet	37
3. Algoritmo ActiveRank	41
3.1. Descripción general	42
3.2. Operaciones básicas	43
3.3. Red de información generada a partir de ActiveRank	48
3.4. Manejo de información utilizando ActiveRank	50
4. Trabajo Experimental	57
4.1. Descripción de sistemas de análisis de información y escenarios analizados	58
4.1.1. Arquitectura de sistemas y escenarios analizados	58
4.1.2. Infraestructura para el procesamiento y almacenamiento de información	60
4.1.3. Función de ActiveRank en el proceso de análisis e indexación	61



Tesis: "Análisis del algoritmo ActiveRank como método de detección automática de contenido dentro de redes de información" Fernando Luege Mateos



México D.F., Febrero 2010

4.2. Metodología de exploración, procesamiento y análisis de información4.3. Resultados obtenidos	62 68
5. Análisis de Resultados	81
5.1. Acerca de la interpretación visual de la matriz de rankings ActiveRank	82
5.2. Proceso de clasificación manual de la muestra	82
5.3. Medición de la eficiencia del proceso de clasificación	84
5.4. Elementos que afectan la eficiencia en el proceso de clasificación	94
5.5. Análisis del dominio <i>unam.mx</i>	96
6. Conclusiones	97
6.1. Conclusiones	98
6.2. Contribuciones	99
6.3. Trabajo futuro	100
7. Apéndices	107
Apéndice A: Dimensiones de la World Wide Web	108
Apéndice B: Diagrama de flujo de un sistema básico de análisis de la WWW	111
Apéndice C: Información estadística sobre la pornografía en Internet	115
Apéndice D: Resultados complementarios	119
8. Referencia Documental	128