

Mesa redonda
El futuro de los buscadores
Nuevas tendencias en recuperación de información

Antonio S. Valderrábanos
Bitext - The Bits and Text Company
www.bitext.com

Fesabid - Santiago, 9-11 de mayo de 2007
Patrocina MasMedios





El planteamiento

- Antes - La biblioteca
- La mayor concentración existente de conocimiento hasta 199X
- Ahora - La web
- La mayor concentración existente de conocimiento pasa a ser una gota en el océano

La biblioteca

Llega internet y la información ubicua

El planteamiento

- Antes - La biblioteca

- El experto mas cualificado es

el bibliotecario
el documentalista

- Ahora - La web

- ¿Dónde está el experto?

Diseñador de buscadores

El planteamiento

¿Hemos ganado o hemos perdido?

- En cuanto a cantidad de información,

hemos ganado

- En cuanto a expertos,

hemos perdido

El planteamiento

Hemos salvado los muebles: los buscadores

- Los buscadores sirven, aunque regular
 - “The inmates are running the asylum”
 - “Trapped in the Net: The Unanticipated Consequences of Computerization”
- Los documentalistas conservan el saber hacer respaldado por siglos de experiencia
- Pero están fuera del juego

El resultado

- Los buscadores sirven, aunque regular
 - No distinguen mayúsculas de minúsculas
 - acentos o diacríticos
 - ... no entienden la materia prima que manejan: el contenido
 - “palabra” = 70 61 6C 61 62 72 61
- Como consecuencia
 - “cana” y “caña”
 - dan el mismo resultado
 - “políticos Sudán”
 - “los políticos también sudan”

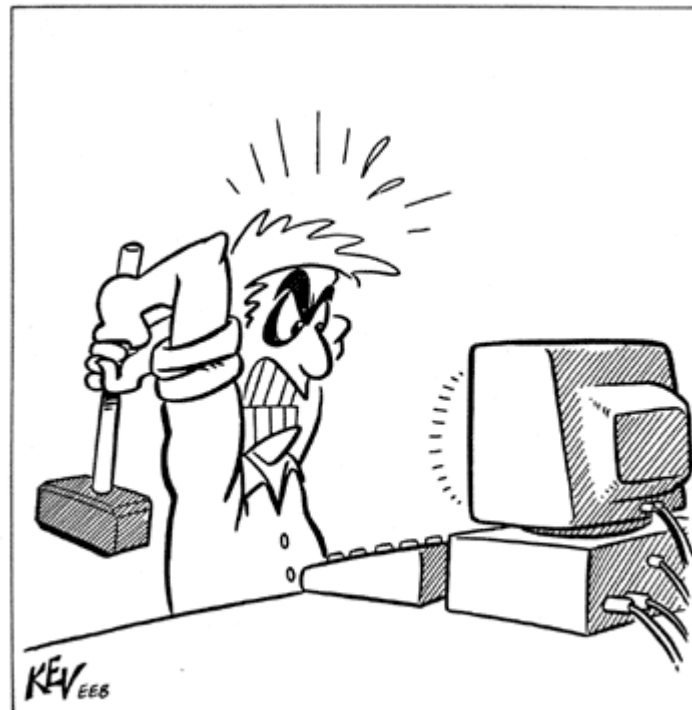
La solución

- Necesitamos que los buscadores
 - conserven la cantidad de información que gestionan
 - recuperar el saber hacer perdido: el documentalista
- Cómo
 - El documentalista tiene que
 - mantener todo su bagaje cinetífico
 - y jugar en el terreno de la informática
- Pero no es así
 - [IWETEL] Discusión sobre Vivissimo

Un ejemplo - Bitext

- El problema
 - Analizar lenguaje natural usando máquinas
 - Procesamiento de lenguaje - Tecnología lingüística
- Dos tendencias
 - Ingenieros: técnicas estadísticas, lógica difusa...
 - Lingüistas: técnicas clásicas (las del colegio)
- Bitext
 - ha llevado las técnicas clásicas al terreno informático
 - las aplica en el entorno de los buscadores
 - porque los buscadores manejan TEXTO, no números

El usuario y los sistemas no se entienden



--PRESS ANY KEY TO CONTINUE--

Cómo lo resolvemos - DataSuite

- **DataGrammar - Intérprete de lenguaje natural**
 - Entiende frases como: “información sobre fusiones entre empresas”
- **DataLexica - Base de datos léxica**
 - Más que un “stemmer” inteligente
- **DataSpell - Corrector ortográfico**
 - Mejora el “Quiso decir...” de Google
- **DataNet - Gestor de relaciones semánticas**
 - Relaciona contenidos automáticamente
- **Características técnicas**
 - Desarrollos propios 100% en C/C++ multiplataforma
 - Fácilmente integrables con cualquier aplicación

NaturalFinder - DataSuite para buscadores

- Soluciona las limitaciones de cualquier buscador
 - Captura todos los documentos relacionados
 - Selecciona los relevantes
 - De manera sencilla y rápida para el usuario
- Sin cambiar de buscador: es integrable con
 - cualquier buscador y en cualquier plataforma
- Y con cualquier fuente de información
 - documentos, bases de datos, híbrido
 - metadatos, tesauros, ontologías
- Implantación sencilla y rápida (residente o remoto)

Conclusión

El futuro de los buscadores
pasa porque sean capaces de “entender” nuestro lenguaje

Interfaz - Para que podamos formular consultas con sencillez
Índice - Para que los resultados sean mejores

Gracias por la atención

Bitext.com – The Bits and Text Company
www.bitext.com – info@bitext.com

Antonio S. Valderrábanos - Fundador y Director
asv@bitext.com

Bitext.com