

Bitext - The Bits and Text Company

NaturalFinder- El complemento ideal para su buscador

Bitext - The Bits and Text Company

+34 911461672

info@bitext.com

<http://www.bitext.com>

¿Cómo funciona un buscador estándar?

- Dada una palabra o palabras
 - fusiones empresariales
 - fusiones empresas
- Localiza ocurrencias de estas palabras
 - ocurrencias exactas
 - “empresariales”, pero no “empresas”
 - “fusiones”, pero no “fusión”
 - “fusiones”, pero no “fusionar”, “fusionadas”...
 - ocurrencias en cualquier unidad de contenido
 - “... empresa italiana... interesada en fusión fría...”

Limitaciones de un buscador estándar

- Una consulta como “fusiones entre empresas”
 - no recupera documentos como
 - “... la fusión entre ambas empresas ...”
 - “... la empresa protagonista de las fusiones ...”
 - “... las empresas fusionadas ...”
 - “... las empresas que se fusionaron ...”
 - a pesar de que son relevantes
- **Esta estrategia genera pérdidas**
 - de tiempo, para usuarios, empleados...
 - de dinero, por pérdida de información, repetición de tareas...

Contenido

Introducción

- ✓ ¿Cómo funciona un buscador estándar?
- ✓ Limitaciones de un buscador estándar

Qué es NaturalFinder

- ✓ Ventajas
- ✓ Características del Servicio
- ✓ Características Técnicas
- ✓ Arquitectura

Análisis lingüístico

Rentabilidad

Ventajas de NaturalFinder

Corrección ortográfica

- Evita consultas sin respuesta: “inmobiliaria”
 - ✓ Detecta problemas de ortografía y tipografía
 - ✓ y propone alternativas similares
- Mejora el “Quiso decir...” de Google
más de 3.000.000 de palabras

Optimización de la consulta

- Detección de palabras relacionadas en la búsqueda
 - ✓ fusiones >> fusión, fusionadas, se fusionaron...
- Mejora los operadores de truncamiento
 - ✓ pago* >> pagos... pagoda
- Incluye relaciones semánticas: sinónimos, hiperónimos...
 - ✓ ventajas >> beneficios
 - ✓ ciudad >> Madrid

Ventajas de NaturalFinder

Interfaz en lenguaje natural

- Adapta el buscador al lenguaje del usuario
“necesito documentación sobre planes de pensiones”

Filtrado por relevancia lingüística: relativa y absoluta

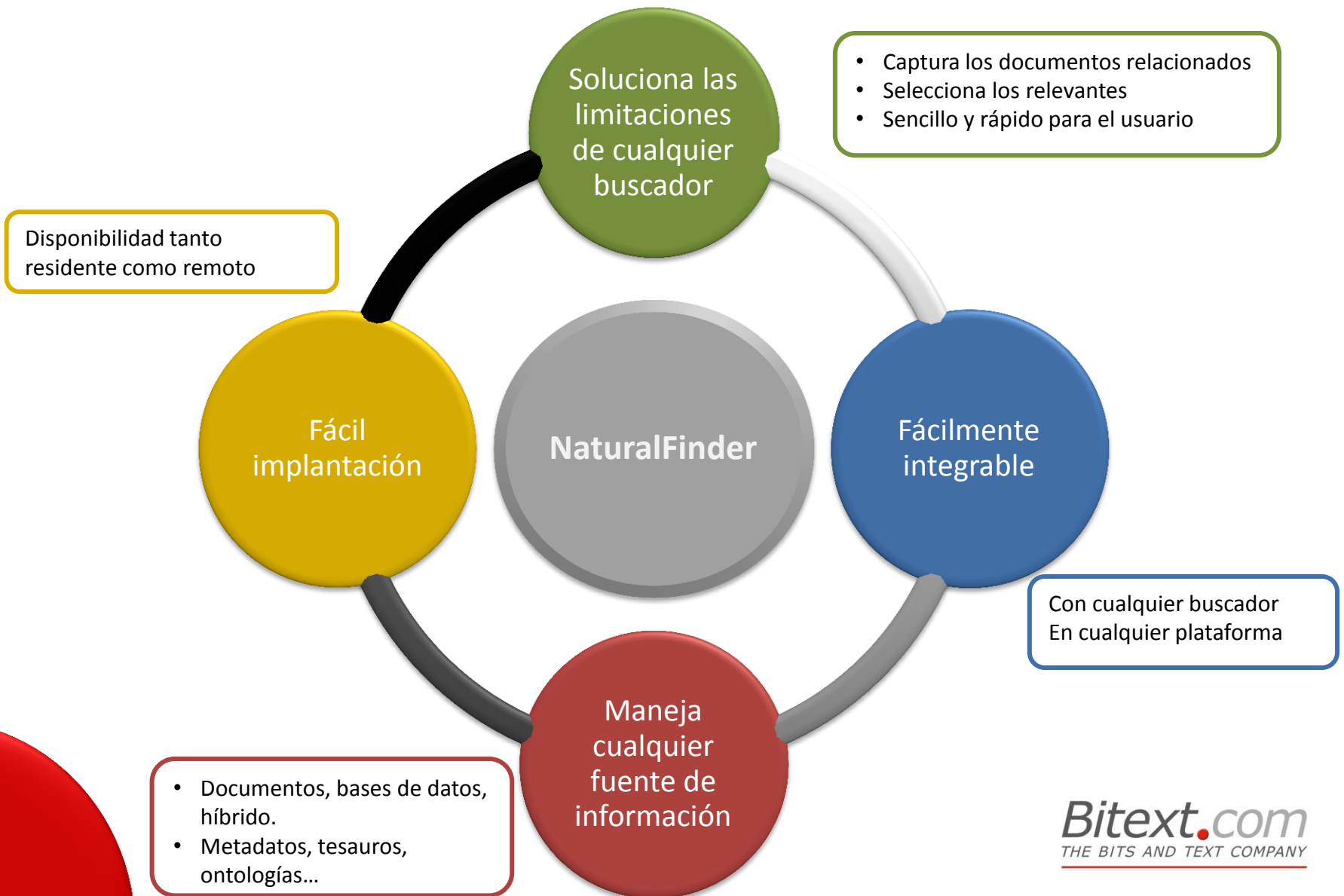
- Los mejores resultados
 - ✓ Cobertura óptima sin añadir ruido
- Con el menor esfuerzo
 - ✓ Hace innecesario el uso de operadores
 - ✓ Pone las opciones avanzadas al alcance de todos

Ventajas de NaturalFinder

Resumen

- Los mejores resultados
- Cobertura óptima sin añadir ruido
- Con el menor esfuerzo
- Hace innecesario el uso de operadores
- Pone las opciones avanzadas al alcance de todos

Características del Servicio de NaturalFinder



Características técnicas de NaturalFinder

Desarrollos

- C/C++ multiplataforma (Windows y Linux/UNIX)
- Componentes modulares (DLL, SO)
- API flexible (C/C++, .NET, Java...)

Fácilmente integrable con cualquier buscador

- dtSearch, Google Search Appliance , Lucene, , Autonomy, Memex...

La salida puede adaptarse a cualquier lenguaje forma

- Operadores booleanos, SQL, expresiones regulares

Características técnicas de NaturalFinder

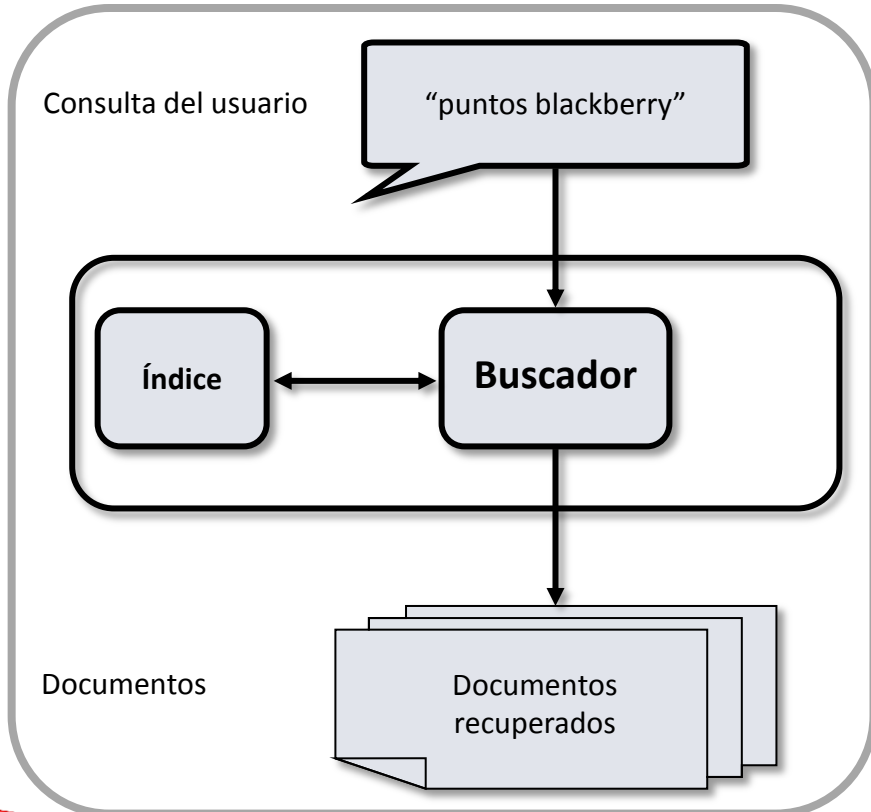
DataSuite: núcleo de NaturalFinder

Necesidades de memoria: entre 1 y 5MB (código y datos)

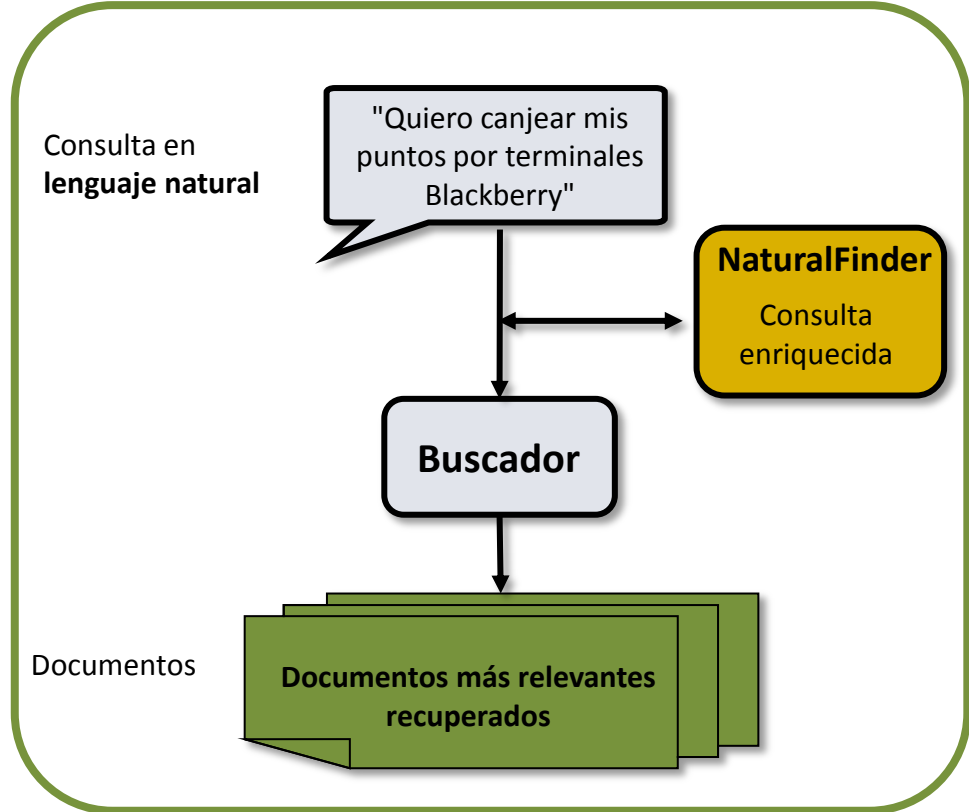
- Nivel léxico
 - ✓ Más de 3.000.000 palabras codificadas lingüísticamente
 - ✓ Procesa cientos de miles de palabras por segundo
- Nivel gramatical:
 - ✓ Cobertura completa del idioma
 - ✓ Procesa miles de frases por segundo
- Idiomas: español, inglés, francés, euskera, catalán y valenciano
 - ✓ En desarrollo en gallego, portugués, italiano y alemán
 - ✓ Adaptable fácilmente a otros idiomas y dominios

Arquitectura

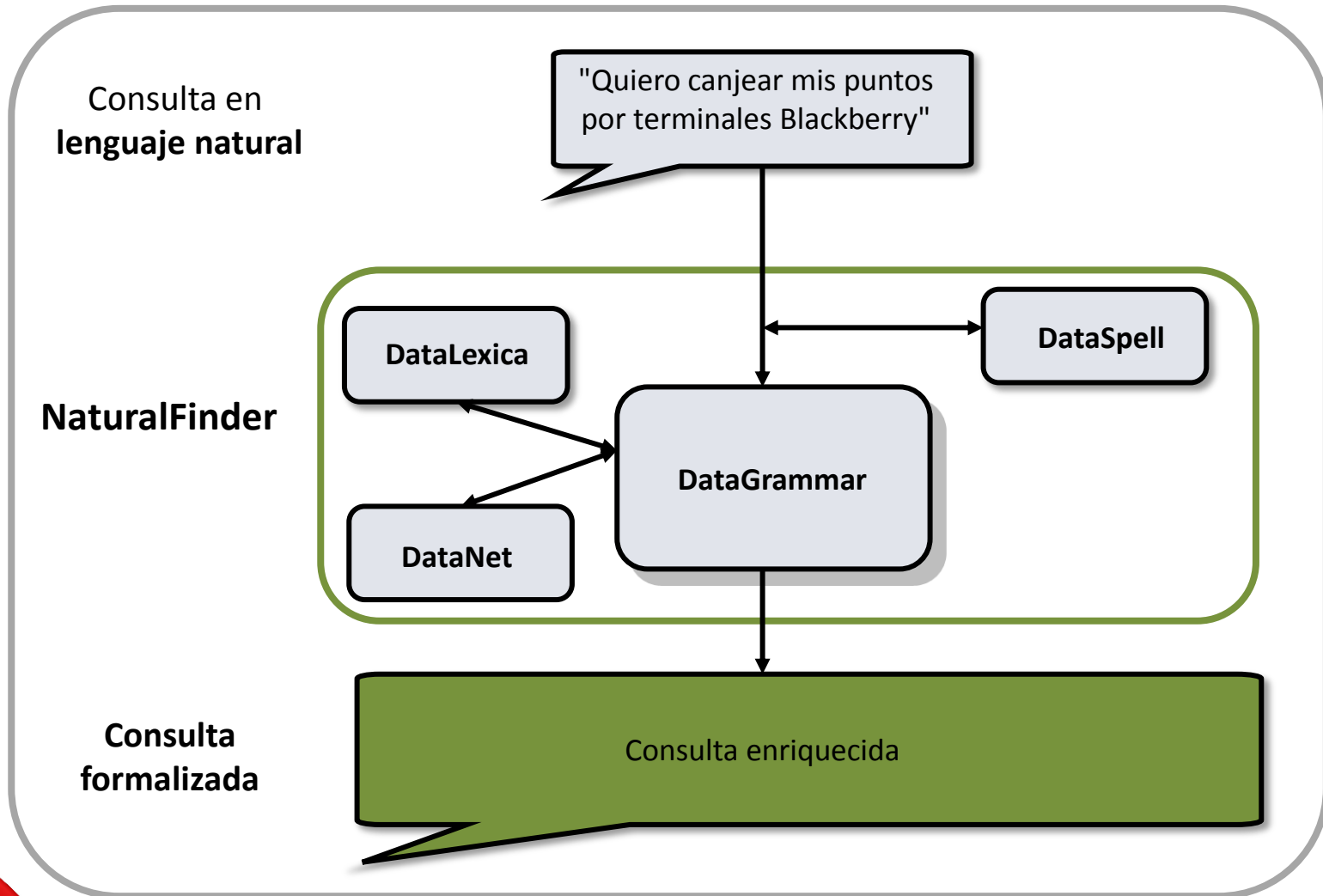
Buscador tradicional



DataSuite + buscador



Arquitectura – Detalle DataSuite



Contenido

Introducción

- ✓ ¿Cómo funciona un buscador estándar?
- ✓ Limitaciones de un buscador estándar

Qué es NaturalFinder

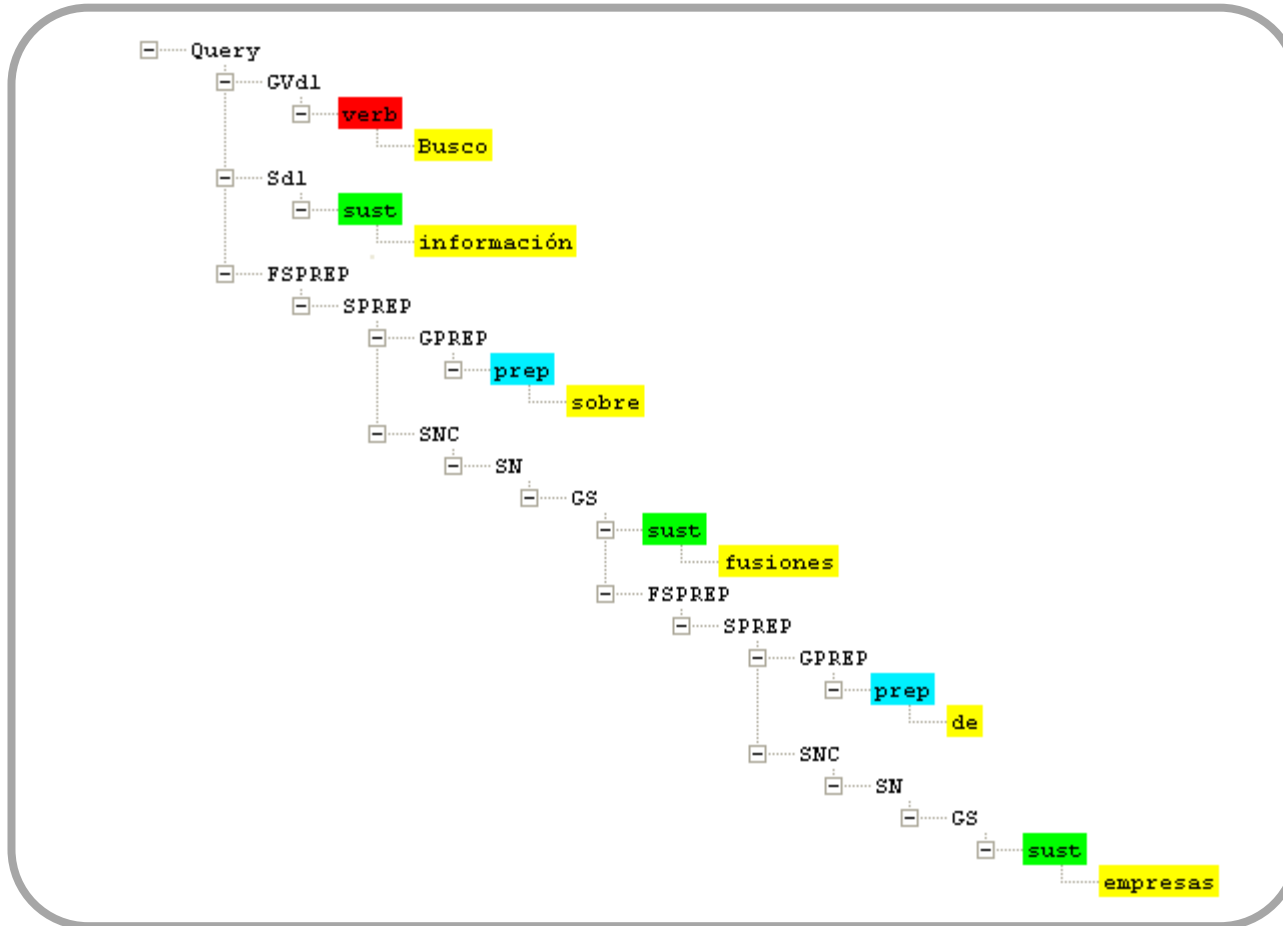
- ✓ Ventajas
- ✓ Características del Servicio
- ✓ Características Técnicas
- ✓ Arquitectura

Análisis lingüístico

Rentabilidad

Análisis lingüístico

“Busco información sobre fusiones de empresas”



Contenido

Introducción

- ✓ ¿Cómo funciona un buscador estándar?
- ✓ Limitaciones de un buscador estándar

Qué es NaturalFinder

- ✓ Ventajas
- ✓ Características del Servicio
- ✓ Características Técnicas
- ✓ Arquitectura

Análisis lingüístico

Rentabilidad

Rentabilidad

- Explotación comercial
 - Coste de la información no encontrada
 - “The High Cost of Not Finding Information”
 - IDC White Paper, Susan Feldman & Chris Sherman
 - Usuarios
 - Usuarios satisfechos
 - Captación de nuevos usuarios (tecnóforos)
 - Pago por descarga (CD/DVD, suscripción vía web)
 - Uso gratuito del buscador (a través de internet)
 - Explotación de contenidos por documento descargado
 - Amplía el potencial de compra
 - Evita “pago por ruido”

Gracias



Bitext - The Bits and Text Company

Edificio Prisma, 1, 1

Calle Cólquide 6

28230 Las Rozas, Madrid

Tel: +34 91 146 1672

info@bitext.com

<http://www.bitext.com>